

66	25	2.0	3767	6	AR127480	AR127480	Sequence
67	26	2.0	3767	6	AR127481	AR127481	Sequence
68	26	2.0	3767	6	AR173300	AR173300	Sequence
69	26	2.0	3767	6	AR173301	AR173301	Sequence
70	26	2.0	3767	6	AF084530	AF084530	Homo sapi
71	26	2.0	3870	3	BT001891	BT001891	Drosophila
72	26	2.0	4382	6	AR243293	AR243293	Sequence
73	26	2.0	4874	6	BD082199	BD082199	Actin
74	26	2.0	6859	3	AF020902	AF020902	Microtub
75	26	2.0	7474	6	A58522	A58522	Sequence
76	26	2.0	9697	6	A31752	A31752	Sequence
77	26	2.0	9697	6	BD085108	BD085108	Vertebra
78	26	2.0	10443	6	A58521	A58521	Sequence
79	26	2.0	110000	6	AC097795	AC097795	Rattus no
80	26	2.0	137001	2	AP001005	AP001005	Homo sapi
81	26	2.0	138115	2	AC142040	AC142040	Rattus no
82	26	2.0	142000	9	AB000845	AB000845	Homo sapi
83	26	2.0	142207	2	BC078890	BC078890	Oryza sat
84	26	2.0	146830	2	EX470168	EX470168	Danio rer
85	26	2.0	161198	2	AC015867	AC015867	Homo sapi
86	26	2.0	163480	2	AC015867	AC015867	Homo sapi
87	26	2.0	172801	10	AC115122	AC115122	Mus muscu
88	26	2.0	173058	2	EX537271	EX537271	Danio rer
89	26	2.0	174882	2	AC104564	AC104564	Homo sapi
90	26	2.0	177352	2	EX005311	EX005311	Danio rer
91	26	2.0	178861	2	AC102743	AC102743	Mus muscu
92	26	2.0	178868	2	AC0668231	AC0668231	Homo sapi
93	26	2.0	180389	2	AL357752	AL357752	Human DNA
94	26	2.0	185102	9	EX470148	EX470148	Danio rer
95	26	2.0	188346	2	AC0113139	AC0113139	Homo sapi
96	26	2.0	190792	5	AC010189	AC010189	Homo sapi
97	26	2.0	213847	2	AL953910	AL953910	Zebrafish
98	26	2.0	215263	2	AC120262	AC120262	Rattus no
99	26	2.0	216598	2	EX0050405	EX0050405	Danio rer
100	26	2.0	218632	2	EX248129	EX248129	Danio rer
					EX248304	EX248304	Danio rer

ALIGNMENTS

RESULT 1				
AX375462				
LOCUS	AX375462	1323 bp	DNA	linear
DEFINITION	Sequence 3 from Patent WO0196370.			
ACCESSION	AX375462			
VERSION	AX375462.1	GI:19170053		
KEYWORDS	.			
SOURCE	Cryptosporidium parvum			
ORGANISM	Cryptosporidium parvum Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Cryptosporidiidae; Cryptosporidium.			
REFERENCE	1			
AUTHORS	Simonsen,J.N.			
TITLE	Cryptosporidium parvum antigens, antibodies thereto and diagnostic and therapeutic compositions thereof			
JOURNAL	Patent: WO 0196370-A 3 20-DEC-2001;			
	University of Manitoba (CA)			
FEATURES	Location/Qualifiers			
source	1..1323			

CDS

BASE COUNT
ORIGIN

481 a	255 c	206 g	381 t
-------	-------	-------	-------

Query Match	100.0%	Score 1323;	DB 6;	Length 1323;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1323;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	GAATTCGGCAGCAATTTTTTTTTTTCTTTTACCTATTTCAATAGTTCTTTGATTCAA	60
OY	1	GAATTCGGCAGCAATTTTTTTTTTTCTTTTACCTATTTCAATAGTTCTTTGATTCAA	60
Db	1	GAATTCGGCAGCAATTTTTTTTTTTCTTTTACCTATTTCAATAGTTCTTTGATTCOA	60
OY	61	ACGATGCAGAAATCTATTTTGTGTTTAAATCCAGATGATCCGGAATTTTGAAAAACATT	120
Db	61	ACGATGCAGAAATCTATTTTGTGTTTAAATCCAGATGATCCGGAATTTTGAAAAACATT	120
OY	121	CTACTAAATTCGAAATTTAAATTTGAGCTTGGCTTGATTTAAATGATAGTTGGCTCGGAGTG	180
Db	121	CTACTAAATTCGAAATTTAAATTTGAGCTTGGCTTGATTTAAATGATAGTTGGCTCGGAGTG	180
OY	181	ATATTTTTATCCTTGATAGGAAACAGCCTCTTGAAGCTGTAAAGTTAATTCATTCGCTTGTG	240
Db	181	ATATTTTTATCCTTGATAGGAAACAGCCTCTTGAAGCTGTAAAGTTAATTCATTCGCTTGTG	240
OY	241	TTTTCTATACAAAACATGTTTGAAGAAAGTAGACAATTGTCTTTAAACCTTTTATC	300
Db	241	TTTTCTATACAAAACATGTTTGAAGAAAGTAGACAATTGTCTTTAAACCTTTTATC	300
OY	301	GGCGTGAGAAATAAATGACTTTTGGTCTGAGAAAGCTTAGGGAACAATCTCCAACTT	360
Db	301	GGCGTGAGAAATAAATGACTTTTGGTCTGAGAAAGCTTAGGGAACAATCTCCAACTT	360
OY	361	CTAATTCGAAATTTATCTTTTCCCTACATGGAATTAATCTGATCTTTGATATATTCOA	420
Db	361	CTAATTCGAAATTTATCTTTTCCCTACATGGAATTAATCTGATCTTTGATATATTCOA	420
OY	421	TTAAACACAAATAAATCAACAAATTTTCTGTTAAAGTTCCGGAGAAACCTCAATCTTCA	480
Db	421	TTAAACACAAATAAATCAACAAATTTTCTGTTAAAGTTCCGGAGAAACCTCAATCTTCA	480
OY	481	CAATTTGGATTTGAGATTAATTCATGATGATTTTCTCTCTTTAATTTAAACATGCGCTA	540
Db	481	CAATTTGGATTTGAGATTAATTCATGATGATTTTCTCTCTTTAATTTAAACATGCGCTA	540
OY	541	TAAAGAAAGAAATGGCCACTTGATATATCAAGATCAAAATATTCAAAAATGAATATAGAG	600
Db	541	TAAAGAAAGAAATGGCCACTTGATATATCAAGATCAAAATATTCAAAAATGAATATAGAG	600
OY	601	CCACTAAATCTAAACAAATCCAGAAATCGAGTGGGAAATTAATTAATCTGACACCAAAACCA	660
Db	601	CCACTAAATCTAAACAAATCCAGAAATCGAGTGGGAAATTAATTAATCTGACACCAAAACCA	660
OY	661	AAGCCTCATCAGTTGTTAGACCGCATCTCTACAGAAAAGCCCTCAAAATGTGGAATCATCA	720
Db	661	AAGCCTCATCAGTTGTTAGACCGCATCTCTACAGAAAAGCCCTCAAAATGTGGAATCATCA	720
OY	721	GAATCTGGTTCAAGACAAGCCCTTAATTAACCTCACAGAAAACGAATCAAGTTCAAATCAT	780
Db	721	GAATCTGGTTCAAGACAAGCCCTTAATTAACCTCACAGAAAACGAATCAAGTTCAAATCAT	780
OY	781	CCTTCTGTGACAGTTCCAGATACATGATACGATTAATTCCTCTGTATTAATTTCCAGA	840
Db	781	CCTTCTGTGACAGTTCCAGATACATGATACGATTAATTCCTCTGTATTAATTTCCAGA	840
OY	841	GACTGATTCAGACTCAGATCACGCGCTTGTGACAAATTCAGAGACTGGATCAGTTCAAA	900
Db	841	GACTGATTCAGACTCAGATCACGCGCTTGTGACAAATTCAGAGACTGGATCAGTTCAAA	900
OY	901	TCATCTTCTGTACTATTAACAGAAACAGGATCCAGCTCAGATCAACTCTGTCTACTCTC	960
Db	901	TCATCTTCTGTACTATTAACAGAAACAGGATCCAGCTCAGATCAACTCTGTCTACTCTC	960
OY	961	CAGAAAGAAAGATTGGATCTCAGAAAGCTTAACCAATCACTTTCAGAAACAAATCAAGCA	1020
Db	961	CAGAAAGAAAGATTGGATCTCAGAAAGCTTAACCAATCACTTTCAGAAACAAATCAAGCA	1020

QY 1021 GCTACATATCTTACCAAGAAAATGAAAATCATATATAGAGAGTAATTCGATTTT 1080
 DB 1021 GCTACATATCTTACCAAGAAAATGAAAATCATATATATAGAGAGTAATTCGATTTT 1080
 QY 1081 AATACATTAATTTTCCAAATGCAACCAATCTTTCAGCAAGTGCAGATGCGAAAGT 1140
 DB 1081 AATACATTAATTTTCCAAATGCAACCAATCTTTCAGCAAGTGCAGATGCGAAAGT 1140
 QY 1141 TATGGGAAAAGATAAAATGTTGATGTCAGCAAGTAATCACTAAAATGACATTAAT 1200
 DB 1141 TATGGGAAAAGATAAAATGTTGATGTCAGCAAGTAATCACTAAAATGACATTAAT 1200
 QY 1201 GAAGATATCTTCCAAAGAAATTTAGAAAATAATGATCTGATGTAATATGCG 1260
 DB 1201 GAAGATATCTTCCAAAGAAATTTAGAAAATAATGATCTGATGTAATATGCG 1260
 QY 1261 CTGAGCATTTCCAAATATTTATTTGTCACCTCAAGTAATAAAAAATTTTAAATC 1320
 DB 1261 CTGAGCATTTCCAAATATTTATTTGTCACCTCAAGTAATAAAAAATTTTAAATC 1320
 QY 1321 GAG 1323
 DB 1321 GAG 1323

RESULT 2
 AF178459 1310 bp mRNA linear INV 01-SEP-2000
 LOCUS AF178459
 DEFINITION Cryptosporidium parvum unknown mRNA sequence.
 ACCESSION AF178459
 VERSION AF178459.1 GI:9957266
 KEYWORDS Cryptosporidium parvum
 SOURCE Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 ORGANISM Cryptosporidium parvum

REFERENCE
 AUTHORS Simonen,J.N., Kramer,M., Lowden,C. and Wilkins,J.
 TITLE Neutralising antigen 2 recognized by human immune sera
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1310)
 AUTHORS Simonen,J.N., Kramer,M., Lowden,C. and Wilkins,J.
 TITLE Direct Submission
 JOURNAL Submitted (16-AUG-1999) Medical Microbiology, University of
 Manitoba, 730 William Ave, Winnipeg, MB R3E 0W3, Canada
 COMMENT NCBI staff are still waiting for submitters to provide appropriate
 feature information.
 FEATURES
 source Location/Qualifiers
 1..1310
 /organism="Cryptosporidium parvum"
 /mol_type="mRNA"
 /db_xref="taxon:5807"

BASE COUNT 470 a 253 c 206 g 381 t

Query Match 61.8%; Score 817; DB 3; Length 1310;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 817; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 501 ATGATGATTTTCTTCTTTAATTAACATGCGCTATAAAGAAAATGGCCACTTG 560
 DB 494 ATGATGATTTTCTTCTTTAATTAACATGCGCTATAAAGAAAATGGCCACTTG 553
 QY 561 AATTAATCAAGATCAAAATTTTCAAAATGAATTAAGAACCCCTAATCAACAAATCC 620
 DB 554 AATTAATCAAGATCAAAATTTTCAAAATGAATTAAGAACCCCTAATCAACAAATCC 613
 QY 621 AGAATCAGTGGGAATTAATTAATCTGAGAACCAAAACCAAGCCTCATCAGTTGTTAG 680
 DB 614 AGAATCAGTGGGAATTAATTAATCTGAGAACCAAAACCAAGCCTCATCAGTTGTTAG 673
 QY 681 ACCGCATCTTACAGAAAAGCCCTCAATGTTGAACATCAAGAAATCTGTTCAAGCAAGC 740

DB 674 ACCGCATCTTACAGAAAAGCCCTCAATGTTGAACATCAAGAAATCTGTTCAAGCAAGC 733
 QY 741 CCTATTACCTCAGCAAGAAAAGAAATCAAGTTCAATCATCTTCTGTGACAGTTCCAG 800
 DB 734 CCTATTACCTCAGCAAGAAAAGAAATCAAGTTCAATCATCTTCTGTGACAGTTCCAG 793
 QY 801 TACTGATAGTTCAAAATCTCTTCTGTGTTACTATTTCAGAGACTGGATCAGACTCAGATC 860
 DB 794 TACTGATAGTTCAAAATCTCTTCTGTGTTACTATTTCAGAGACTGGATCAGACTCAGATC 853
 QY 861 AGCGCCTTGTCAGCAATTCAGAGACTGGATCAGTTCAATCATCTTCTGTGTTACTAC 920
 DB 854 AGCGCCTTGTCAGCAATTCAGAGACTGGATCAGTTCAATCATCTTCTGTGTTACTAC 913
 QY 921 AGAAAAGAGATCAGCTCAGATCAGACTCAGCTGCTACTTCTCCAGAAAGAGATTGACTCA 980
 DB 914 AGAAAAGAGATCAGCTCAGATCAGACTCAGCTGCTACTTCTCCAGAAAGAGATTGACTCA 973
 QY 981 GAACGTTACCAATCACTTCTTACAGAAACAACCTCAAGCAGCTACATATCTTACCAAGA 1040
 DB 974 GAACGTTACCAATCACTTCTTACAGAAACAACCTCAAGCAGCTACATATCTTAAACCAAG 1033
 QY 1041 AATGAAATCATTAATTAATCAGAAAGTAATTCGAGTTTAAATACATAATCTTCCAAA 1100
 DB 1034 AATGAAATCATTAATTAATCAGAAAGTAATTCGAGTTTAAATACATAATCTTCCAAA 1093
 QY 1101 TCAACCAATCTTTCACGCAAGCTGCAGATGTGAAAGTTATGGGAAAAGATTAAT 1160
 DB 1094 TCAACCAATCTTTCACGCAAGCTGCAGATGTGAAAGTTATGGGAAAAGATTAAT 1153
 QY 1161 GTTGATGTTGAGCAAGTAATCACTTAATAATGACATTTATGAAGATCTTCCAGAAAGAT 1220
 DB 1154 GTTGATGTTGAGCAAGTAATCACTTAATAATGACATTTATGAAGATCTTCCAGAAAGAT 1213
 QY 1221 TGAAGCAAAATGTAAGATCTGATGTAATTAATGCGCTTATGACATTTCCAAATATC 1280
 DB 1214 TGAAGCAAAATGTAAGATCTGATGTAATTAATGCGCTTATGACATTTCCAAATATC 1273
 QY 1281 TAAATGTCACTCAAGTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 1317
 DB 1274 TAAATGTCACTCAAGTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 1310

RESULT 3
 MEU79565 1520 bp mRNA linear MAM 05-JAN-1999
 LOCUS MEU79565
 DEFINITION Macropus eugenii RBM1 mRNA, complete cds.
 ACCESSION U79565
 VERSION U79565.1 GI:4098579
 KEYWORDS Macropus eugenii (tammar wallaby)
 SOURCE Macropus eugenii
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euleleostomi;
 Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.

REFERENCE
 AUTHORS Delbridge,M.L., Harry,J.L., Toder,R., O'Neill,R.J.W., Ma,K.,
 TITLE A human candidate spermatogenesis gene RBM1 is conserved and
 JOURNAL amplified on the marsupial Y chromosome
 REFERENCE 2 (bases 1 to 1520)
 AUTHORS Delbridge,M.L. and Graves,J.A.M.
 TITLE Direct Submission
 JOURNAL Submitted (25-NOV-1996) Genetics and Human Variation, La Trobe
 University, Kingsbury Drive, Bundoora, Victoria 3083, Australia
 FEATURES
 source Location/Qualifiers
 1..1520
 /organism="Macropus eugenii"
 /mol_type="mRNA"
 /db_xref="taxon:9315"
 /chromosome="Y"
 62..1270
 /note="RNA-binding motif gene"

CDS

Query Match	Best Local Similarity	100.0%; Pred. No. 0.0022;	Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
<p>BASE COUNT 475 a 265 c 397 g 383 t</p> <p>ORIGIN</p>	<p>2.2%; Score 29; DB 4; Length 1520;</p> <p>100.0%; Pred. No. 0.0022;</p> <p>Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>	<p>1295 AAGTAAACCTCGAG 1323</p> <p>1492 AAGTAAACCTCGAG 1520</p>	<p>/codon_start=1</p> <p>/product="RBM1"</p> <p>/protein_id="AAD00327.1"</p> <p>/db_xref="GI:4098580"</p> <p>/translation="MEACPGTFLIGLAVGTNEKDELSVEFKYGHIVKVLKMDQET</p> <p>NKSGFLVHVFESPAADKADAMNGKALDCKSIKEVDANPFESGRRPPAPRS</p> <p>GPPGLRVRGSGSGARPPRGKMDSDGYGLNFMGSSMGPLLVKRGA PQSGQSP</p> <p>KRSPGSGVYSSROMGGRGOLSGGRDYGCPRRSELSRRDYMSPPDSDSKES</p> <p>SSRYSASSMDGRGYASFRPDYAPLPREYTRDYGHSSSDYTSKRYSDRSDYGG</p> <p>RDRYSDHSGRGYRDSYESGYRRAPPPGSPCYGSSSYEDYSTRDYGCGCG</p> <p>SQKYSRSSRGIVYSSRDCVGRORGLPSPMSRGYPAPHDYSSSSRGARGDGRGS</p> <p>RSERGGRSY"</p>
<p>RESULT 4</p> <p>LOCUS AF069994 1525 bp mRNA linear VRT 28-DEC-1998</p> <p>DEFINITION Danio rerio vimentin mRNA, complete cds.</p> <p>ACCESSION AF069994</p> <p>VERSION AF069994.1 GI:4063848</p> <p>KEYWORDS</p> <p>SOURCE Danio rerio (zebrafish)</p> <p>ORGANISM Danio rerio</p> <p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;</p> <p>Actinopterygii; Neopterygii; Teleostei; Ostariophysi;</p> <p>Cypriniformes; Cyprinidae; Danio.</p> <p>1 (bases 1 to 1525)</p> <p>Cerda,J., Conrad,M., Markl,J., Brand,M. and Herrmann,H.</p> <p>Direct Submission</p> <p>Submitted (03-JUN-1998) Div. Cell Biology, German Cancer Research</p> <p>Center, Im Neuenheimer Feld 280, Heidelberg 69120, Germany</p> <p>FEATURES</p> <p>source</p>	<p>1. 1525</p> <p>/organism="Danio rerio"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:7955"</p> <p>/dev_stage="postmitogenesis embryo"</p> <p>48. 1415</p> <p>/function="cytoskeletal protein"</p> <p>/note="Intermediate filament; similar to Swiss-Prot</p> <p>Accession Number P48673 and to product encoded by Genbank</p> <p>Accession Number S76850"</p> <p>/codon_start=1</p> <p>/product="vimentin"</p> <p>/protein_id="AAC98491.1"</p> <p>/db_xref="GI:4063849"</p> <p>/translation="MASRTSYSKRMFGGERPAARSTFSSROYSPVRSSTSY</p> <p>NYSAPSPVYSSKGLRVSGAPLPRLATIDLPGLAEINTEFRANTRNEAKEMQHLN</p> <p>DRPSYIDKVRFLKQONKILAELEKAGSAGVDLDEEMELRRVQLGEKSGS</p> <p>KVEVDNLIGEDIRLEKERLOEMLQREDAENSLSRPROVDNLSARLDERVESL</p> <p>OEELAPLKULHDELAEMQIOMQOHQIODEYVTKPDLTALKRVYQYETLARINQ</p> <p>ESBEWYKSPMADSEEAARNEALRLAQENDRYROLQALCDLELAKGTNESLEKQ</p> <p>MRELEWYSSMASGYODTIGRLDEDIRNMKDEMARHREYODLLNFKALDIEATYR</p> <p>KILGEGESRITSPPNPFSSSLRSMKETRPLDNLISKVVIKTIETRDGHI INESTQ</p> <p>NEDLE"</p>	<p>BASE COUNT 439 a 382 c 422 g 282 t</p> <p>ORIGIN</p>	<p>2.1%; Score 28; DB 5; Length 1525;</p> <p>100.0%; Pred. No. 0.0072;</p> <p>Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>

```

Db      1498 ACTAAAAAAAAAAAAAAAAACTCGAG 1525

RESULT 5
AL512586/c 65975 bp DNA linear ROD 26-MAR-2003
LOCUS      Mouse DNA sequence from clone RP23-37017 on chromosome 4, complete
DEFINITION
ACCESSION  AL512586
VERSION     AL512586.12
KEYWORDS   GI:27764025
SOURCE     HTG.
ORGANISM   Mus musculus (house mouse)
MUS musculus
Eukaryote, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Scuriognathia, Muridae, Murine, Mus.
1 (bases 1 to 65975)
North, P., Leaves, N., Greystron, J., Coppola, M., Manjunath, S.,
Russell, E., Smith, M., Strachan, G., Tofts, C., Boal, E., Cobley, V.,
Hunter, G., Kimberley, C., Thomas, D., Cave-Berry, L., Weston, P. and
Bocherdy, M. R. M.
Direct Submission
Submitted (03-FEB-2003) Mouse Sequencing Group, HGMP-RC, Hinxton,
Cambridge, CB10 1SB, UK. E-mail enquiries:- mbochoche@hmp.mrc.ac.uk
or pnoir@hmp.mrc.ac.uk
HGMP-RC part of the UK Mouse Sequencing Consortium
On Jan 15, 2003 this sequence version replaced gi:21530910.
----- Genome Center
Center: UK Medical Research Council
Center code: UK-MRC
Web site: http://mrcseq.har.mrc.ac.uk
Contact: mouse@har.mrc.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a VGC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL, Sw, SWISSPROT, Tr, TREMBL, Wp, WORMPEP, Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep/ RP23-37017 is
from the RPCT-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBac3.6
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
IMPORTANT: This sequence is not the entire insert of clone
RP23-37017 It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP23-37017 is at 1 in this sequence.
Location/Qualifiers
1..65975
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-37017"
/clone_1bp="RPCT-23"
repeat_region 1..1256

```



```

repeat_region /note="L1_MM repeat: matches 5319. .6571 of consensus"
1291. .1312
/note="22.0 copies 1 mer A 35% conserved"
repeat_region complement(1326. .1470)
/note="MER109 repeat: matches 870. .1018 of consensus"
repeat_region complement(1497. .1969)
/note="MER109 repeat: matches 216. .1325 of consensus"
repeat_region /note="MER109 repeat: matches 216. .1325 of consensus"
2000. .2096
/note="Lx repeat: matches 6876. .6972 of consensus"
repeat_region 2214. .2224
/note="2.8 copies 4 mer ATTT 22% conserved"
repeat_region 2286. .2306
/note="2.1 copies 10 mer CTCATGAT 33% conserved"
repeat_region 2300. .2309
/note="2.5 copies 4 mer TATG 20% conserved"
repeat_region 2621. .2638
/note="4.5 copies 4 mer TTAA 22% conserved"
repeat_region 2625. .2636
/note="2.0 copies 6 mer TTAAT 24% conserved"
repeat_region 3748. .3763
/note="2.0 copies 8 mer TATACATG 32% conserved"
repeat_region 3781. .3804
/note="12.0 copies 2 mer AT 21% conserved"
repeat_region 3786. .3804
/note="2.7 copies 7 mer TATATAC 31% conserved"
repeat_region 4684. .4696
/note="2.6 copies 5 mer TTTC 26% conserved"
repeat_region 4904. .5047
/note="B1_MM repeat: matches 1. .144 of consensus"
repeat_region 5048. .5074
/note="6.8 copies 4 mer CAAA 47% conserved"
repeat_region 5169. .5458
/note="B4A repeat: matches 2. .288 of consensus"
repeat_region 5529. .5538
/note="2.5 copies 4 mer CATG 20% conserved"
repeat_region 5549. .5601
/note="2.9 copies 18 mer AGAAGTAAGAGGAGG 56% conserved"
repeat_region 5558. .5569
/note="3.0 copies 4 mer GAGG 24% conserved"
repeat_region 5592. .5617
/note="6.5 copies 4 mer GAGG 27% conserved"
repeat_region 5596. .5610
/note="2.1 copies 7 mer GAGGGA 30% conserved"
repeat_region 5615. .5631
/note="2.1 copies 8 mer GGAAGAA 34% conserved"
repeat_region 5779. .5924
/note="B1_MM repeat: matches 1. .144 of consensus"
repeat_region 5925. .5948
/note="6.0 copies 4 mer CAAA 39% conserved"
repeat_region complement(5947. .6458)
/note="match: GSS: Em:A2067786"
repeat_region complement(6249. .6515)
/note="ORR1D repeat: matches 98. .374 of consensus"
repeat_region 6516. .6814
/note="L1_MM repeat: matches 6272. .6569 of consensus"
repeat_region complement(6883. .6987)
/note="ORR1D repeat: matches 1. .104 of consensus"
repeat_region 7095. .7107
/note="13.0 copies 1 mer A 26% conserved"
repeat_region 7216. .7227
/note="2.4 copies 5 mer TTATT 24% conserved"
repeat_region 7489. .7509
/note="1.9 copies 11 mer AAAACATTCA 42% conserved"
repeat_region complement(7534. .7648)
/note="PB1D9 repeat: matches 9. .118 of consensus"
repeat_region 7662. .7674
/note="3.2 copies 4 mer TCTG 26% conserved"
repeat_region complement(7714. .8619)
/note="Lx9 repeat: matches 6637. .7549 of consensus"
repeat_region complement(8624. .10220)
/note="Lx9 repeat: matches 4987. .6587 of consensus"
repeat_region 10221. .10238
/note="18.0 copies 1 mer T 27% conserved"

```

```

repeat_region complement(10239. .11934)
repeat_region /note="L1_MM repeat: matches 4895. .6588 of consensus"
1291. .1312
/note="22.0 copies 1 mer A 35% conserved"
repeat_region complement(11906. .12933)
/note="Lx9 repeat: matches 6122. .6646 of consensus"
12283. .12304
/note="Lx9 repeat: matches 6122. .6646 of consensus"
misc_feature /note="Lx9 repeat: matches 6122. .6646 of consensus"
repeat_region complement(12412. .12546)
/note="Lx9 repeat: matches 5837. .5975 of consensus"
repeat_region 12669. .12994
/note="ORR1B1 repeat: matches 36. .365 of consensus"
repeat_region 12998. .13022
/note="12.5 copies 2 mer TC 23% conserved"
repeat_region 13041. .13059
/note="1.9 copies 10 mer TCATACCTCA 29% conserved"
repeat_region 13062. .13090
/note="4.8 copies 6 mer CCTCT 35% conserved"
repeat_region 13063. .13080
/note="3.6 copies 5 mer CCTCT 27% conserved"
repeat_region 13537. .14258
/note="match: GSS: Em:A2844194"
repeat_region 14602. .14614
/note="2.2 copies 6 mer TCTACT 26% conserved"
repeat_region 15540. .15707
/note="B2_MM2 repeat: matches 1. .173 of consensus"
repeat_region complement(15708. .16251)
/note="Lx2 repeat: matches 5345. .5890 of consensus"
repeat_region 16249. .17234
/note="Lx2 repeat: matches 5884. .6850 of consensus"
repeat_region 17319. .17329
/note="2.2 copies 5 mer TCTG 22% conserved"
repeat_region 17337. .17350
/note="2.8 copies 5 mer TTTC 28% conserved"
repeat_region 18307. .18316
/note="2.5 copies 4 mer TTTA 20% conserved"
repeat_region 18518. .18536
/note="2.1 copies 9 mer AATTTCG 38% conserved"
repeat_region 18751. .18770
/note="2.5 copies 8 mer TGTTCG 33% conserved"
repeat_region 18896. .18907
/note="2.0 copies 6 mer ATGCAT 24% conserved"
repeat_region complement(19023. .19284)
/note="ORR1D repeat: matches 68. .368 of consensus"
repeat_region 19285. .19327
/note="21.5 copies 2 mer AG 86% conserved"
repeat_region complement(19328. .19496)
/note="ORR1D repeat: matches 1. .181 of consensus"
repeat_region 19497. .19551
/note="27.5 copies 2 mer CA 92% conserved"
repeat_region 19880. .19891
/note="2.0 copies 6 mer TACTTT 24% conserved"
repeat_region 19902. .19922
/note="2.1 copies 10 mer TTTCCATA 33% conserved"
repeat_region 20580. .20589
/note="2.5 copies 4 mer AGAC 20% conserved"
repeat_region 20833. .20845

```

Query Match 2.1% Score 28; DB 10; Length 65975;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1290 AACTCAAGTAAAAA 1317
|||||
Db 60146 AACTCAAGTAAAAA 60119
|||||

RESULT 6
BD056965/c 300 bp DNA linear PAT 27-AUG-2002

DEFINITION Sets of labeled energy transfer fluorescent primers and their use in multi component analysis.

ACCESSION BD056965

VERSION BD056965.1 GI:22602571

KEYWORDS JP 2001509271-A/2.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 300)

U,J.

Set of labeled energy transfer fluorescent primers and their use in multi component analysis

Patent: JP 2001509271-A 2 10-JUL-2001;

INCYTE PHARMACEUTICALS INC

PN JP 2001509271-A/2

PD 10-JUL-2001

PF 12-DEC-1997 JP 1998534358

PR 15-JAN-1997 JP 08/784162

PI JINGYUE JU

PC G01N21/78, C12N15/09, C12Q1/68, C12N15/00

CC Strandedness: Single;

CC Topology: Linear;

FH Key Location/Qualifiers.

1.300 Location/Qualifiers

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

BASE COUNT 58 a 79 c 81 g 80 t 2 others

ORIGIN

Query Match 2.0%; Score 27; DB 6; Length 300;

Best Local Similarity 100.0%; Pred. No. 0.027;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAACTCGAG 1323

Db 85 GTAAAAAAAAAAAAAAAACTCGAG 59

RESULT 7

LOCUS G72232/c 640 bp mRNA linear STS 21-DEC-2001

DEFINITION Ssn10 stickleback cDNA Gasterosteus aculeatus STS cDNA clone

ACCESSION G72232

VERSION G72232.1 GI:17977092

KEYWORDS STS.

SOURCE Gasterosteus aculeatus (three spined stickleback)

ORGANISM Gasterosteus aculeatus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.

1 (bases 1 to 640)

Pichel,C.L., Nereng,K.S., Ohgi,K.A., Cole,B.L., Colosimo,P.F., Buerkle,C.A., Schluter,D. and Kingsley,D.M.

The genetic architecture of divergence between threespine stickleback species

Nature 414 (6866), 901-905 (2001)

11780061

SYNOPSIS: C3greens

Contact: David M. Kingsley

Dept of Developmental Biology/ Howard Hughes Medical Institute

Stanford University School of Medicine

8300 Beckman Center, 279 Campus Drive, Stanford, CA 94305-5329, USA

Email: kingsley@cmm.stanford.edu

Primer A: AATGACCTGACAGAGGCC

Primer B: CTTGACCGACCTGTGATCC

STS size: 106

PCR Profile:

Denaturation: 95 degrees C for 1:45

Annealing: 56 degrees C for 0:45

Polymerization: 72 degrees C for 0:45

PCR cycles: 1

Denaturation: 94 degrees C for 0:45

Annealing: 56 degrees C for 0:45

Polymerization: 72 degrees C for 0:45

PCR cycles: 4

Denaturation: 90 degrees C for 0:45

Annealing: 56 degrees C for 0:45

Polymerization: 72 degrees C for 0:45

PCR cycles: 30

Thermal cycler: MJ Research PTC-200 DNA Engine

Protocol:

Template: 5 ng

Primer: each 0.5 uM

dNTPs: each 62.5 uM

Taq polymerase: 0.25 units

Total volume: 10 ul

Buffer:

MgCl2: 1.5 mM

KCl: 50 mM

Tris-HCl: 10 mM

pH: 8.3.

1.640 Location/Qualifiers

/organism="Gasterosteus aculeatus"

/mol_type="mRNA"

/strain="Salinas River, California"

/db_xref="taxon:69293"

/clone="55-22/17"

/tissue type="Head and internal organs"

/clone.lib="Stickleback cDNA"

/dev_stage="adult"

/lab_host="X11-Blue MR"

/note="Vector: ZAP Express; V-type: Phage; Oligo(dT) primed cDNA unidirectionally cloned into the EcoRI/XhoI sites of ZAP Express."

45.150

45.64

STS primer_bind 45.150

primer_bind 121 c 114 g 181 t 15 others

BASE COUNT 209 a 121 c 114 g 181 t 15 others

ORIGIN

Query Match 2.0%; Score 27; DB 11; Length 640;

Best Local Similarity 100.0%; Pred. No. 0.026;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAACTCGAG 1323

Db 27 GTAAAAAAAAAAAAAAAACTCGAG 1

RESULT 8

LOCUS A62439 1017 bp DNA linear PAT 09-MAR-1998

DEFINITION Sequence 1 from Patent WO9713853.

ACCESSION A62439

VERSION A62439.1 GI:3716359

KEYWORDS

SOURCE Aspergillus niger

ORGANISM Aspergillus niger

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

REFERENCE 1

AUTHORS Van,D.B., De,G.L., Visser,J. and Van,O.A.

TITLE PROTEIN DETECTION

JOURNAL Patent: WO 9713853-A 1 17-APR-1997;

GIST BROCADES BV (NL)

Other publication AU 7294396 970430.

COMMENT

FEATURES

source 1.1017

CDS

/organism="Aspergillus niger"
/mol_type="genomic DNA"
/strain="N400"
/isolate="CBS120.49"
/db_xref="taxon:5061"
57..776

/codon_start=1
/product="CELLULOSE"
/protein_id="CAA03652.1"
/db_xref="GI:3716360"
/translation="MKLPVSLAHLAATAMGQTCGSDASPPYSVNONLMEGYCT
GSCGVVDKLSGSGASGHTWTGSGEGTYSNSGVTNNKLVSDVSIPISVYRK
QDNTNADVAIDLFTANVDHATSSGDYELMLVLRVGIQIGQIATATVGGKSW
EWMVSGTQAGAQRTYSFVSESPINSYSGDINAFPSYLTONGFPASSOYLINLQFG
TEAFTCGPAFTVDNMTASVN"

BASE COUNT

258 a 267 c 265 g 227 t

Query Match 2.0%; Score 27; DB 6; Length 1017;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1323
Db 991 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1017

RESULT 9

LOCUS A62449 1017 bp DNA linear PAT 09-MAR-1998
DEFINITION Sequence 1 from Patent WO9713862.
ACCESSION A62449
VERSION A62449.1 GI:3716369
KEYWORDS
SOURCE Aspergillus niger
ORGANISM Aspergillus niger
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

REFERENCE 1
AUTHORS Van D.B., De G.L., Visser, J. and Van O.A.
TITLE FUNGAL CELLULASES
JOURNAL Patent: WO 9713862-A 1 17-APR-1997;
GIST BROCADES BV (NL)
COMMENT Other publication AU 7294496 970430.
FEATURES
source location/Qualifiers

1..1017

/organism="Aspergillus niger"
/mol_type="genomic DNA"
/strain="N400"
/isolate="CBS120.49"
/db_xref="taxon:5061"
57..776
/codon_start=1
/product="CELLULOSE"
/protein_id="CAA03657.1"
/db_xref="GI:3716370"
/translation="MKLPVSLAHLAATAMGQTCGSDASPPYSVNONLMEGYCT
GSCGVVDKLSGSGASGHTWTGSGEGTYSNSGVTNNKLVSDVSIPISVYRK
QDNTNADVAIDLFTANVDHATSSGDYELMLVLRVGIQIGQIATATVGGKSW
EWMVSGTQAGAQRTYSFVSESPINSYSGDINAFPSYLTONGFPASSOYLINLQFG
TEAFTCGPAFTVDNMTASVN"

CDS

BASE COUNT

258 a 267 c 265 g 227 t

Query Match 2.0%; Score 27; DB 6; Length 1017;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1323
Db 991 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1017

RESULT 10

LOCUS AR130937 1017 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6190890.
ACCESSION AR130937
VERSION AR130937.1 GI:14119262
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1017)

AUTHORS Van Den Broeck, H.C., De Graaff, L.H., Visser, J. and Van
Ooijen, A.J.J.
TITLE Fungal cellulases

JOURNAL Patent: US 6190890-A 1 20-FEB-2001;
VERSION Location/Qualifiers

FEATURES

1..1017

BASE COUNT 258 a 267 c 265 g 227 t

Query Match 2.0%; Score 27; DB 6; Length 1017;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1323
Db 991 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1017

RESULT 11

LOCUS AR174096 1017 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6306635.
ACCESSION AR174096
VERSION AR174096.1 GI:17914416
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1017)

AUTHORS Van Den Broeck, H.C., De Graaff, L.H., Visser, J. and Van
Ooijen, A.J.J.
TITLE Fungal cellulases

JOURNAL Patent: US 6306635-A 1 23-OCT-2001;
VERSION Location/Qualifiers

FEATURES

1..1017

BASE COUNT 258 a 267 c 265 g 227 t

Query Match 2.0%; Score 27; DB 6; Length 1017;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1323
Db 991 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1017

RESULT 12

LOCUS AY075210 1045 bp mRNA linear INV 01-FEB-2002
DEFINITION Drosophila melanogaster AT14709 full length cDNA.
ACCESSION AY075210
VERSION AY075210.1 GI:18446971
KEYWORDS FLI_CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1045)

[illegible]

CDS

/db_xref="taxon:3930"
/clone="CIGPDH30"
/clone_11b="C-DNA-BANK ZAP"
34..1152
/note="unnamed protein product"
/codon_start=1
/protein_id="CA02755.1"
/db_xref="GI:2298897"
/translation="MAPSELCTHONPHSSGYDGRSRVTVGSGNMGSVAALIASN
TKLPSPHDEVMVPEETLPSGEKLTLDVNNNNVKTLPGLGRNVVADDDLENA
VKRANMLVFETPHQPMEGICRKLVGKIQEQAQSLIKMEYCKECPMTSSISDL
GNGCVLMGANIANEIAVEKFSATVGFRENTDIAEKVQLESTPIFMVAEDVGC
ELCGTKNIVAIAAGVFDGLEKGNNTKAAIMRIGLRKMAFSKLPSPVADTFEFG
GVADLTTCGGRNRKVAEAPAKGGERSEFDLEAEILRGOKIQGVSTAKEVEYVGH
RGWLPLPFPSTHISTGRHPSAIVESEQRTFSM"

polya_signal 1349..1354 356 g 343 t
BASE COUNT 400 a 291 c 356 g 343 t
ORIGIN

Query Match 2.0%; Score 27; DB 6; Length 1390;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAATCTCGAG 1323
Db 1364 GTAAAAAATCTCGAG 1390

RESULT 16
ARI06141 1390 bp DNA linear PAT 14-FEB-2001
LOCUS ARI06141 7 from patent US 6103520.
DEFINITION ARI06141
ACCESSION ARI06141.1 GI:12820206
VERSION
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1390)
AUTHORS Topfer,R., Hausmann,L. and Schell,J.
TITLE Plant DNA encoding glycerol-3-phosphate dehydrogenase (GPDH)
JOURNAL Patent: US 6103520-A 7 15-AUG-2000;
FEATURES
source 1..1390
BASE COUNT 400 a 291 c 356 g 343 t
ORIGIN

Query Match 2.0%; Score 27; DB 6; Length 1390;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAATCTCGAG 1323
Db 1364 GTAAAAAATCTCGAG 1390

RESULT 17
PL1293506 3841 bp mRNA linear INV 06-JAN-2001
LOCUS PL1293506
DEFINITION Paracentrotus lividus mRNA for kinesin-like protein boursein
(boursein gene).
ACCESSION AJ293506
VERSION AJ293506.1 GI:12055368
KEYWORDS boursein; boursein gene; kinesin-like.
SOURCE Paracentrotus lividus (common urchin).
ORGANISM Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Echinoidea; Echinacea; Echinoida; Echinidae;
Paracentrotus.

REFERENCE 1
Toulou,I., Lhomond,G. and Puillere,G.
AUTHORS Boursein, a sea urchin b1mc motor protein, plays a role in anaphase

and cytokinesis
Unpublished
2 (bases 1 to 3841)
REFERENCE Toulou,I.
AUTHORS
JOURNAL Direct Submission
Submitted (10-JUL-2000) Toulou I., CNRS UMR 7009, Laboratoire de
Biologie du developpement, Observatoire Oceanologique, 06230
Villefranche s/mer, FRANCE
LOCATION/Qualifiers
1..3841
/organism="Paracentrotus lividus"
/mol_type="mRNA"
/db_xref="taxon:7656"
1..3841
/gene="boursein"
109..3354
/gene="boursein"
/function="role in mitosis and cytokinesis"
/product="kinesin-like boursein"
/codon_start=1
/db_xref="GI:12055369"
/translation="MKKLRPPQEKIQQVAVCRPSSEKKNSYSVDVPRAKREI
IVTEVAEKASSKTFSDKVFGEKSTQIVSVVAPILDEVLMGINTVFAYGQGT
GRTFTWEGDRTDPLSWEQDPLAGIIPRPMQIPEKMGTVDFSVRSYLEYNE
LPDLGQEDTQMRIFEDSARKGSYVIGLEEVTVHNNKVEYVAILKSGASKQTAT
LNNASRSRSVSSTIHIKENSIDDELKTGKLVNVLAGESENIGSGAVDRARE
AGNINOSLITLGRVITAVEHA PHVPRRSKLTJRILODLSGRTKSIATVSPASIN
VEETSLTDYARAKKITRPEINOKTLKALIKETTERIBERKDLPTRENGIFL
SEHYRSMETSIASQAQIKEMENIEGLTQMOKTELFEYQKLEBORTLELTIT
KOLVETDITLHTEKDLRTTQDRDQORLVEHVTVTEQLMSAEATQVSTDSVD
VGLSKLDRKKNVEAHNKSAQEVFAESFRSHSDIKSALQREKQSCASMOQOF
EEMISKRTKEADLTALSDMTVSVGOTGALMEBQREKREKWSADGACBSKFK
EDVIDRIDPHENRPLTMTKLTLEKLSVTELSNERKDKISQSTOSVOCIDKO
LEOVKMTTVKRPADHOTSRTTDLSGQDELKEHRRHNMKSIQDLFSRESF
ASETSKLDQRFREASTDTATVADLOSOTAIQESCESTRVHKDSERLHSSADHH
KFTVGVEDIRKYNASLSESVSHNTLLERQTHYQCEKVKGMDEVLQVITQOE
ADTEAMTNEHLASSETLSTTVDKOLEMDCTVGMTEKSKTSSELTTEEMGME
NGLKREVEFTLQDLKEDVPTGTPQKQVSYPRDLSTEPHDIILNRFQECCLNAA
DMPLPDDSVASCPADISLQSGKTQSSFLSDTSVDVDAENETSSISRSQOOSLSS
KPFADDTKENVMPKSVTTARGKLPDRSRKRTPTKTSRLPRANSVYN"

BASE COUNT 1201 a 849 c 920 g 871 t
ORIGIN

Query Match 2.0%; Score 27; DB 3; Length 3841;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAATCTCGAG 1323
Db 3815 GTAAAAAATCTCGAG 3841

RESULT 18
BX321868 135734 bp DNA linear HTG 02-JUN-2003
LOCUS BX321868
DEFINITION Danio rerio clone RP71-41121. *** SEQUENCING IN PROGRESS ***
unorderd pieces.
ACCESSION BX321868
VERSION BX321868.3 GI:30424196
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 135734)
McLaren,S.
AUTHORS Direct Submission
JOURNAL Submitted (28-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-helpsanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On May 7, 2003 this sequence version replaced gi:29500978.

```
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: b241121
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 135414 bases at least Q40
Consensus quality: 135526 bases at least Q30
Consensus quality: 135589 bases at least Q20
Insert size: 135634; sum-of-coverage
Insert size: 143080; 4.1% error; agarose-fp
Quality coverage: 7.35x in Q20 bases; sum-of-coverage
Quality coverage: 6.97x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 106542: contig of 106542 bp in length
* 106543 106642: gap of 100 bp
* 106643 135734: contig of 23092 bp in length.
Location/Qualifiers
1. 135734
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="RP71-41121"
/clone_id="RPCT-71"
1. 106542
/feature="assembly_fragment:00302"
fragment chain:1
clone end:T7
vector_side:left"
106643. 135734
/feature="assembly_fragment:00153"
fragment chain:1
clone end:SP6
vector_side:right"
misc_feature
BASE COUNT 46602 a 23697 c 23799 g 41336 t 100 others
ORIGIN
Query Match 2.0%; Score 27; DB 2; Length 135734;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1292 CTCACGTAAGTAAAAAAAAAAAAAAAAAC 1318
|||||
Db 91065 CTCACGTAAGTAAAAAAAAAAAAAAAAAC 91091
|||||

RESULT 19
BX088528 169422 bp DNA linear HTG 03-MAY-2003
LOCUS Danio rerio clone CH211-226D23, *** SEQUENCING IN PROGRESS ***, 5
DEFINITION Unordered pieces.
ACCESSION BX088528
VERSION BX088528.4 GI:30349910
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
CYPRINIFORMES; Cyprinidae; Danio.
Pelam,S.
REFERENCE 1 (bases 1 to 169422)
AUTHORS
```

```
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
COMMENT zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 3, 2003 this sequence version replaced gi:28971433.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zc226D23
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 168238 bases at least Q40
Consensus quality: 168618 bases at least Q30
Consensus quality: 168833 bases at least Q20
Insert size: 169022; sum-of-coverage
Insert size: 159347; 7.6% error; agarose-fp
Quality coverage: 5.66x in Q20 bases; sum-of-coverage
Quality coverage: 6.00x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4434: contig of 4434 bp in length
* 4435 4534: gap of 100 bp
* 4535 45363: contig of 40829 bp in length
* 45364 45463: gap of 100 bp
* 45464 55734: contig of 50271 bp in length
* 55735 95834: gap of 100 bp
* 95835 101616: contig of 5782 bp in length
* 101617 101716: gap of 100 bp
* 101717 169422: contig of 67706 bp in length.
Location/Qualifiers
1. 169422
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-226D23"
/clone_id="CHOR1-211"
1. 4434
/feature="assembly_fragment:00375"
fragment chain:1"
4535. 45363
/feature="assembly_fragment:00876"
fragment chain:1"
45464. 55734
/feature="assembly_fragment:00410"
fragment chain:1"
95835. 101616
/feature="assembly_fragment:00299"
101717. 169422
/feature="assembly_fragment:01281"
misc_feature
BASE COUNT 53118 a 29749 c 31377 g 54778 t 400 others
ORIGIN
Query Match 2.0%; Score 27; DB 2; Length 169422;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1290 AACTCAAGTAAAAAAAAAAAAAAAAAAAA 1316
|||||
Db 9400 AACTCAAGTAAAAAAAAAAAAAAAAAAAA 9374
|||||

RESULT 20
```

AC136114 179844 bp DNA linear HTG 23-NOV-2002
 LOCUS Rattus norvegicus clone CH230-482C23, *** SEQUENCING IN PROGRESS
 REFERENCE 1 (bases 1 to 179844)
 AUTHORS Muzny, D., Marle, M., Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsebrook, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biwalto, K., Blair, J., Blankenship, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Cree, A., D'Souza, C., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, C., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Dlyva, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C., Gabler, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodge, A., Hogue, M., Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kover, C., Kowalski, C., Kraft, C., L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensen, L., Louie, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaelele, O., Okunolu, G., Olarnpunasoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., River, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., D., Smalls, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valae, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D., R., Holt, R., A., Smith, H., O., Weinrock, G., and Gibbs, R., A.

TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 179844)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (29-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 179844)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission

JOURNAL Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Nov 23, 2002 this sequence version replaced gi:2417985. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KCCK
 Center clone name: CH230-482C23
 ----- Summary Statistics
 Assembly program: Phrap, version 0.990329
 Consensus quality: 169377 bases at least Q40
 Consensus quality: 170723 bases at least Q30
 Consensus quality: 171556 bases at least Q20
 Estimated insert size: 175244; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 178114: contig of 178114 bp in length
 * 178115 178214: gap of unknown length
 * 178215 179844: contig of 1630 bp in length.
 Location/Qualifiers
 1. 179844
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-482C23"
 2355..3242
 misc_feature
 /note="clone boundary
 /clone_end:17
 site:MboI
 end sequence: B2133876"
 144786..146133
 misc_feature
 /note="wgs contig"
 174100..174979
 misc_feature
 /note="clone boundary
 /clone_end:5p6
 site:MboI
 end sequence: B2133877"
 176181..178114
 misc_feature
 /note="wgs_end_extensions
 /clone_end:5p6"

BASE COUNT 54378 a 35679 c 34450 g 47277 t 8060 others
 ORIGIN
 Query Match 2.0%; Score 27; DB 2; Length 179844;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 27; Conservative 0; Mismatches 0; Gaps 0;

QY 1291 ACTCACTAAAAA 1317
 Db 46843 ACTCACTAAAAA 46869
 RESULT 21
 AC099128
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-61017, *** SEQUENCING IN PROGRESS
 *** 3 unordered pieces.
 AC099128 219064 bp DNA linear HTG 10-MAY-2003
 AC099128.5 GI:30522425
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 219064)
 Muzny,D,Marie, Metzker,M,Lee, Abramson,S, Adams,C, Alder,J,
 Allen,C, Allen,H, Alibrooks,S, Amin,A, Angiano,D,
 Ayala-Bechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
 Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
 Biswal,K, Blair,D, Blankenburg,K, Blyth,P, Brown,M,
 Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
 Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A,
 Chacko,J, Chavez,D, Chen,G, Chen,Y, Chen,Z, Chu,J,
 Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
 David,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
 Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Diya,K,
 Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
 Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
 Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,T, Garza,P,
 Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Gatzert,
 Gebregeorgis,E, Geier,K, Gill,R, Grady,M, Guerra,W, Guevara,W,
 Gunaratne,P, Haaland,W, Hamli,C, Hamilton,C, Hamilton,K,
 Harvey,Y, Havlik,P, Hawes,A, Henderson,N, Hernandez,J,
 Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,
 Hollins,B, Howells,S, Hulik,S, Hume,J, Idelbird,D, Jackson,A,
 Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
 Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Koval,C,
 Kowitz,C, Kraft,C,L, Lebow,H, Levan,T, Lewis,L, Li,Z, Liu,J,
 Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
 Lorenshuwa,L, Louised,H, Lozano,R,J, Lu,X, Ma,J,
 Meshwari,M, Mahindartine,M, Mahmoud,M, Mallory,K, Mangum,A,
 Mangum,B, Mapa,P, Martin,K, Martin,R, Martinez,E,
 Mawhinney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,
 Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
 Morgan,M, Morris,K, Morris,S, Muidasa,M, Murphy,M, Nairt,L,
 Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,
 Nwokwelenh,O, Okunonu,G, Olarunpasegon,A, Pal,S, Parks,K,
 Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C,
 Plopper,F, Polidexter,A, Popovic,D, Pflum,E, Pu,L,L,
 Piazto,M, Quirroz,J, Rachlin,E, Reeves,K, Regier,M,A, Relgh,R,
 Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
 Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,
 Sanders,W, Savary,G, Scherer,S, Scott,G, Shatsman,S, Shen,H,
 Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D,
 Sneed,A, Sodergren,E, Song,X-Z, Sorelle,R, Sosa,J,
 Steimle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C,
 Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K,
 Valas,R, Vera,V, Villanasa,D, Waldron,L, Walker,B, Wang,J,
 Wang,O, Wang,S, Warren,J, Warren,R, Wei,X, White,F,
 Williams,G, Willson,R, Wleczky,R, Wooden,H, Worley,K,
 Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V,
 Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von
 Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O,
 Weinstock,G, and Gibbs,R,A.
 TITLE Unpublished
 JOURNAL Direct Submission
 REFERENCE 2 (bases 1 to 219064)
 AUTHORS Worley,K,C.
 TITLE Direct Submission

JOURNAL Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 219064)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 COMMENT On May 10, 2003 this sequence version replaced gi:23096551.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GGRV
 Center clone name: CH230-61017
 ----- Summary Statistics
 Assembly program: Atlas 3.0/
 Consensus quality: 208005 bases at least Q40
 Consensus quality: 210612 bases at least Q30
 Consensus quality: 212428 bases at least Q20
 Estimated insert size: 223725; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 215640: contig of 215640 bp in length
 * 215641 215740: gap of unknown length
 * 215741 217188: contig of 1448 bp in length
 * 217189 217288: gap of unknown length
 * 217289 219064: contig of 1776 bp in length.
 Location/Qualifiers
 1. 219064
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-61017"
 /size="2688..3653
 /note="clone boundary
 clone end:Sp6
 site:Ecort
 end sequence: BH286065"
 misc_feature
 BASE COUNT 66324 a 39681 c 39623 g 67748 t 5688 others
 ORIGIN
 Query Match 2.0%; Score 27; DB 2; Length 219064;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 27; Conservative 0; Mismatches 0; Gaps 0;
 QY 1291 ACTCACTAAAAA 1317
 |||

Db 144151 ACTCAGTAAAAA 144177

RESULT 22
BX294188
LOCUS
DEFINITION
Accession
BX294188.4 GI:30424245
HTG: HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS
Dario rerio (zebrafish)
SOURCE
Dario rerio
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 263203)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
McLaren, S.
Submitted (28-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk
On May 7, 2003 this sequence version replaced gi:29500486.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zk255G14
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 262031 bases at least Q40
Consensus quality: 262355 bases at least Q30
Consensus quality: 262903; sum-of-contigs
Insert size: 252892; 4.3% error; agarose-IP
Quality coverage: 7.55x in Q20 bases; sum-of-contigs Quality
coverage: 8.22x in Q20 bases; agarose-IP

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 131807: contig of 131807 bp in length
* 131808 131907: gap of 100 bp
* 131908 136826: contig of 4919 bp in length
* 136827 136926: gap of 100 bp
* 136927 189749: contig of 52823 bp in length
* 189750 189849: gap of 100 bp
* 189850 263203: contig of 73354 bp in length.
Location/Qualifiers
1. 263203
/organism="Dario rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_dkey="255G14"
/clone_1bp="DanioKey"
1. 131807
/note="assembly fragment: 00754
fragment_chain:1"
131908 136826
/note="assembly fragment: 01388
fragment_chain:1"
136927 189749
/note="assembly fragment: 01629.0"
189850 263203
/note="assembly fragment: 02002"

BASE COUNT 86165 a 45855 c 45403 g 85476 t 304 others

ORIGIN

Query Match 2.0% Score 27; DB 2; Length 263203;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119473 ACTCAGTAAAAA 119499

RESULT 23
AC104404
LOCUS
DEFINITION
Accession
AC104404.6 GI:30580766
HTG: HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 288526)

REFERENCE
AUTHORS
Muzny, D., Marie, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Aml, A., Anguiano, D.,
Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, N., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Bunay, C., Butcher, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Casar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C., Falis, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P.,
Frazer, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S., Hodge, J., Hodge, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, R., Kovar, C.,
Kowis, C., Kraft, C., Labow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensbuewa, L., Louised, H., Lozano, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindarne, M., Mahoud, M., Mallory, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M., McNeill, T., Meenan, B.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Muntana, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackelam, O., Okunolu, G., Olarnpusagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Plankoch, C.,
Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
Sanders, W., Savary, G., Scherer, S., Scott, C., Sitter, C., Smit, D.,
Shetty, J., Shvartbeyn, A., Sison, I., Sitter, C., Smit, D.,
Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanal, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Wilson, R., Wlezyk, R., Wodden, H., Wotley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 288526)
Worley, K.C.
Direct Submission
Submitted (09-DEC-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 288526)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23268721.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: CH230-77120
Center clone name: CH230-77120
----- Summary Statistics
Assembly program: Atlas 3.0:
Consensus quality: 265077 bases at least Q40
Consensus quality: 267765 bases at least Q30
Consensus quality: 269757 bases at least Q20
Estimated insert size: 280425; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 3951: contig of 3951 bp in length
* 3952 4051: gap of unknown length
* 4052 272586: contig of 268535 bp in length
* 272587 272686: gap of unknown length
* 272687 273801: contig of 1115 bp in length
* 273802 273901: gap of unknown length
* 273902 274923: contig of 1021 bp in length
* 274923 275022: gap of unknown length
* 275023 276196: contig of 1174 bp in length
* 276197 276296: gap of unknown length
* 276297 277777: contig of 1481 bp in length
* 277778 277877: gap of unknown length
* 277878 281014: contig of 3137 bp in length
* 281015 281114: gap of unknown length
* 281115 284238: contig of 3124 bp in length
* 284239 284339: gap of unknown length
* 284339 288526: contig of 4188 bp in length.

FEATURES
source
Location/Qualifiers
1..288526
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-77120"
1..1241
/note="wgs contig"
4052..7376
/note="wgs contig"
BASE COUNT 85180 a 54676 c 54095 g 77047 t 17528 others
ORIGIN

Query Match 2.0%; Score 27; DB 2; Length 288526;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1291 ACTCAAGTAAAAAAAAAAAAAAAAAAAAA 1317
Db 21413 ACTCAAGTAAAAAAAAAAAAAAAAAAAAA 214139

RESULT 24
LOCUS AR225416 317 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 17 from patent US 6444425.
ACCESSION AR225416
VERSION AR225416.1 GI:27263362
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
1 (bases 1 to 317)
AUTHORS Reed, S.G., Lodes, M.J., Mohamath, R. and Secrist, H.
TITLE Compounds for therapy and diagnosis of lung cancer and methods for their use
JOURNAL Patent: US 6444425-A 17 03-SEP-2002;
FEATURES Location/Qualifiers
source 1..317
BASE COUNT 111 a 55 c 78 g 73 t
ORIGIN

Query Match 2.0%; Score 26; DB 6; Length 317;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
Db 292 TAAAAAAAAAAAAAAAAAACTCGAG 317

RESULT 25
LOCUS AX321486 317 bp DNA linear PAT 15-DEC-2001
DEFINITION Sequence 17 from Patent WO0172295.
ACCESSION AX321486
VERSION AX321486.1 GI:17905529
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Reed, S.G., Lodes, M.J., Mohamath, R., Secrist, H., Benson, D.R., Indira, C.Y., Henderson, R.A., Fling, S.P., Aigate, P.A., Elliot, M., Mannion, J. and Kalos, M.D.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0172295-A 17 04-OCT-2001;
FEATURES CORIXA CORPORATION (US)
source Location/Qualifiers
1..317

```

BASE COUNT      111 a      55 c      78 g      73 t
ORIGIN
Query Match      2.0%; Score 26; DB 6; Length 317;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1298 TAAAAAATCTGAG 1323
Db      292 TAAAAAATCTGAG 317

RESULT 26
AP154637/c      425 bp      mRNA      linear      PLN 10-FEB-2003
LOCUS      AP154637
DEFINITION      Nicotiana tabacum clone PR12 mRNA sequence.
ACCESSION      AP154637
VERSION      AP154637.1 GI:8547137
KEYWORDS
SOURCE      Nicotiana tabacum (common tobacco)
ORGANISM      Nicotiana tabacum
REFERENCE      1 (bases 1 to 425)
AUTHORS      Wang, J., Sheehan, M., Brookman, H. and Timko, M.P.
TITLE      Characterization of cDNAs differentially expressed in roots of
JOURNAL      tobacco (Nicotiana tabacum cv Burley 21) during the early stages of
PUBMED      alkaoid biosynthesis
10996241      Plant Sci. 158 (1-2), 19-32 (2000)
REFERENCE      2 (bases 1 to 425)
AUTHORS      Sheehan, M.J., Wang, J. and Timko, M.P.
TITLE      Direct Submision
JOURNAL      Submitted (27-MAY-1999) Biology, University of Virginia, 044 Gilmer
FEATURES
source
1. 425
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/cultivar="Burley 21"
/db_xref="taxon:4097"
/clone="PR12"
/issue_type="root"
/notes="similar to inorganic phosphate transporter"
BASE COUNT      148 a      75 c      51 g      151 t
ORIGIN
Query Match      2.0%; Score 26; DB 8; Length 425;
Best Local Similarity 100.0%; Pred. No. 0.089;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1298 TAAAAAATCTGAG 1323
Db      51 TAAAAAATCTGAG 26

RESULT 27
G72249      520 bp      mRNA      linear      STS 21-DEC-2001
LOCUS      G72249/c
DEFINITION      Stn53 stickleback cDNA Gasterosteus aculeatus STS cDNA clone
ACCESSION      G72249
VERSION      G72249.1 GI:17977109
KEYWORDS
SOURCE      Gasterosteus aculeatus (three spined stickleback)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;

```

```

REFERENCE      Gasterosteidae; Gasterosteus.
AUTHORS      1 (bases 1 to 520)
TITLE      Peichel, C.L., Nereux, K.S., Ohgi, K.A., Cole, B.L., Colosimo, P.F.,
JOURNAL      Buerkle, C.A., Schluter, D. and Kingsley, D.M.
PUBMED      The genetic architecture of divergence between threespine
11780061      stickleback species
Nature 414 (6866), 901-905 (2001)
COMMENT      Synonyms: B7yellow11
Contact: David M. Kingsley
Dept of Developmental Biology/ Howard Hughes Medical Institute
Stanford University School of Medicine
8300 Beckman Center, 279 Campus Drive, Stanford, CA 94305-5329, USA
Email: kingsley@cgm.stanford.edu
Primer A: CTGGGAGATGATGACGCG
Primer B: CTGCACAGATGAAACGG
STS size: 134
PCR Profile:
Denaturation: 95 degrees C for 1:45
Annealing: 56 degrees C for 0:45
Polymerization: 72 degrees C for 0:45
PCR cycles: 1
Denaturation: 94 degrees C for 0:45
Annealing: 56 degrees C for 0:45
Polymerization: 72 degrees C for 0:45
PCR cycles: 4
Denaturation: 90 degrees C for 0:45
Annealing: 56 degrees C for 0:45
Polymerization: 72 degrees C for 0:45
PCR cycles: 30
Thermal cycler: MJ Research PTC-200 DNA Engine
Protocol:
Template: 5 ng
Primer: each 0.5 uM
dNTPs: each 62.5 uM
Taq polymerase: 0.25 units
Total volume: 10 uL
Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3.
Location/Qualifiers
1. 520
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas River, California"
/db_xref="taxon:69293"
/clone="68-26/T7"
/issue_type="Head and internal organs"
/clone_lib="Stickleback cDNA"
/dev_stage="Adult"
/lab_host="XLI-Blue MRP"
/notes="Vector: ZAP Express; V-type: Phage; Oligo(dT)
primed cDNA unidirectionally cloned into the EcoRI/XhoI
sites of ZAP Express."
324. 457
324. 343
primer_bind 324. 343
primer_bind 151 a
BASE COUNT      151 a      127 c      101 g      139 t      2 others
ORIGIN
Query Match      2.0%; Score 26; DB 11; Length 520;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1298 TAAAAAATCTGAG 1323
Db      26 TAAAAAATCTGAG 1

RESULT 28

```

G32092/c 588 bp DNA linear STS 27-OCT-1998
LOCUS G32092
DEFINITION w57 Chicken RPMACroojmans Gallus gallus STS genomic, sequence
tagged site.
ACCESSION G32092
VERSION G32092.1 GI:2138038
KEYWORDS STS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianidae; Gallus.
1 (bases 1 to 588)
Croojmans,R.P., Dijkhof,R.J., van der Poel,J.J. and Groenen,M.A.
New microsatellite markers in chicken optimized for automated
fluorescent genotyping
JOURNAL Anim. Genet. 28 (6), 427-437 (1997)
MEDLINE 98251489
PUBMED 9589584
COMMENT
Contact: Richard Croojmans
Department of Animal Breeding
Wageningen Agricultural University
Box 338, 6700 AH Wageningen, The Netherlands
Email: richard.croojmans@ag.vf.wau.nl
Primer A: AGGTCTCCCAATTGTCTC
Primer B: CTACTCCAGATCTACTC
STS size: 124
PCR Profile:
Presoak: 94 degrees C for 5 minute (s) :
Denaturation: 94 degrees C for 30 seconds
Annealing: 55 degrees C for 45 seconds
Polymerization: 72 degrees C for 90 seconds
PCR Cycles: 35
Thermal Cycler: MJ research
Protocol:
Template: 10-100 ng
Primer: each 295 nM
dNTPs: each 200 uM
Polymerase Goldstar: 0.4 u
Total Vol: 12 kl
Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
Tetramethylammoniumchloride(TMAC): 1 mM
pH: 8.3.
Location/Qualifiers
1..588
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/clone_lib="Chicken RPMACroojmans"
STS
primer_bind 88..107
primer_bind complement(192..211) 200 t 4 others
BASE COUNT 176 a 116 c 92 g 200 t
ORIGIN
Query Match 2.0%; Score 26; DB 11; Length 588;
Best Local Similarity 100.0%; Pred.No.0.086;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 26 TAAAAAAAAAAAAAAAACTCGAG 1
RESULT 29 599 bp mRNA linear STS 21-DEC-2001
LOCUS G72252
DEFINITION Sns56 stickleback cDNA Gasterosteus aculeatus STS cDNA clone

48-20/T7, sequence tagged site.
ACCESSION G72252
VERSION G72252.1 GI:1797712
KEYWORDS STS.
SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 599)
Peichel,C.L., Nereng,K.S., Ohgi,K.A., Cole,B.L., Colosimo,P.F.,
Buerkle,C.A., Schluter,D., and Kingsley,D.M.
The genetic architecture of divergence between threespine
stickleback species
JOURNAL Nature 414 (6866), 901-905 (2001)
PUBMED 11780061
COMMENT
Synonyms: F12green9
Contact: David M. Kingsley
Dept of Developmental Biology/ Howard Hughes Medical Institute
Stanford University School of Medicine
8300 Beckman Center, 279 Campus Drive, Stanford, CA 94305-5329, USA
Email: kingsley@cgm.stanford.edu
Primer A: ATGCTGTGTGTCTGCTGCTC
Primer B: TCATCAAGTGTGTCTAGGG
STS size: 235
PCR Profile:
Denaturation: 95 degrees C for 1:45
Annealing: 56 degrees C for 0:45
Polymerization: 72 degrees C for 0:45
PCR cycles: 1
Denaturation: 94 degrees C for 0:45
Annealing: 56 degrees C for 0:45
Polymerization: 72 degrees C for 0:45
PCR cycles: 4
Denaturation: 90 degrees C for 0:45
Annealing: 56 degrees C for 0:45
Polymerization: 72 degrees C for 0:45
PCR cycles: 30
Thermal cycler: MJ Research PTC-200 DNA Engine
Protocol:
Template: 5 ng
Primer: each 0.5 uM
dNTPs: each 62.5 uM
Tag polymerase: 0.25 units
Total volume: 10 ul
Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3.
Location/Qualifiers
1..599
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas River, California"
/db_xref="taxon:69293"
/clone="48-20/T7"
/tissue_type="Head and internal organs"
/clone_lib="Stickleback cDNA"
/dev stage="Adult"
/lab host="XJ1-Blue MRF"
/note="Vector: ZAP Express: V-type; Phage: Oligo(dT)
primed cDNA unidirectionally cloned into the EcoRI/XhoI
sites of ZAP Express."
STS
primer_bind 278..512
primer_bind complement(493..512)
BASE COUNT 196 a 164 c 90 g 139 t 10 others
ORIGIN
Query Match 2.0%; Score 26; DB 11; Length 599;

Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323

Db 26 TAAAAAAAAAAAAAAAACTCGAG 1

RESULT 30
G72270/c 600 bp mRNA linear STS 21-DEC-2001

LOCUS G72270 600 bp mRNA linear STS 21-DEC-2001

DEFINITION Scn9g Stickleback cDNA Gasterosteus aculeatus STS cDNA clone

ACCESSION G72270 71-16/T7, sequence tagged site.

VERSION G72270.1 GI:17977130

KEYWORDS STS.

SOURCE Gasterosteus aculeatus (three spined stickleback)

ORGANISM Gasterosteus aculeatus

REFERENCE 1 (bases 1 to 600)

AUTHORS Peichel,C.L., Nereng,K.S., Ohgi,K.A., Cole,B.L., Colosimo,P.F.,

TITLE Buerkle,C.A., Schluter,D., and Kingsley,D.M.

JOURNAL The genetic architecture of divergence between threespine

PUBMED stickleback species

COMMENT Nature 414 (6866), 901-905 (2001)

SYNOPSIS 11780061

CONTACT: David M. Kingsley

DEPT of Developmental Biology/ Howard Hughes Medical Institute

Stanford University School of Medicine

8300 Beckman Center, 279 Campus Drive, Stanford, CA 94305-5329, USA

Email: kingsley@cmmg.stanford.edu

Primer A: CTACCTCAACACACACCC

Primer B: TCCTCTTACGACACACAC

STS size: 168

PCR Profile:

Denaturation: 95 degrees C for 1:45

Annealing: 56 degrees C for 0:45

Polymerization: 72 degrees C for 0:45

PCR cycles: 1

Annealing: 94 degrees C for 0:45

Polymerization: 72 degrees C for 0:45

PCR cycles: 4

Annealing: 90 degrees C for 0:45

Polymerization: 56 degrees C for 0:45

PCR cycles: 72 degrees C for 0:45

Thermal cycler: MJ Research PTC-200 DNA Engine

Protocol:

Template: 5 ng

Primer: each 0.5 uM

dNTPs: each 62.5 uM

Tag polymerase: 0.25 units

Total volume: 10 uL

Buffer:

MgCl2: 1.5 mM

KCl: 50 mM

Tris-HCl: 10 mM

pH: 8.3.

Location/Qualifiers

1. 600

/organism="Gasterosteus aculeatus"

/mol_type="mRNA"

/strain="Salinas River, California"

/db_xref="taxon:69293"

/clone="71-16/T7"

/tissue_type="Head and internal organs"

/clone_id="Stickleback cDNA"

/dev_stage="Adult"

/lab_host="XLI-Blue MRF"

/note="Vector: ZAP Express; V-type: Phage; Oligo(dT)

primed cDNA unidirectionally cloned into the EcoRI/XhoI

sites of ZAP Express."

STS

primer_bind 289..456

primer_bind 289..456

BASE COUNT 180 a 121 c 134 g 160 t 5 others

ORIGIN

Query Match 2.0%; Score 26; DB 11; Length 600;

Best Local Similarity 100.0%; Pred. No. 0.086;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323

Db 26 TAAAAAAAAAAAAAAAACTCGAG 1

RESULT 31

LOCUS BC052930 657 bp mRNA linear ROD 02-JUN-2003

DEFINITION Mus musculus cDNA clone IMAGE:1430210, partial cds.

ACCESSION BC052930

VERSION BC052930.1 GI:31324955

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 657)

AUTHORS Strassberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Bucof,R.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Soares,M.B., Bonaldi,M.F., Casavant,T.L.,

Shapiro,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L.,

Sheets,T.E., Brownstein,M.J., Uddin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loggellano,N.A., Peters,G.J.,

Adams,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,

Mckernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huiyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Boutard,G.G., Blakeley,R.W., Touchman,D.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schultz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalske,U., Smaltus,D.E.,

Schneerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 657)

AUTHORS Strassberg,R.

TITLE Direct Submission

JOURNAL Submitted (30-MAY-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: http://mgc.nci.nih.gov

CONTACT: MGC help desk

Email: cgabbs@femail.nih.gov

Tissue Procurement: Marcella Bento Soares, Ph.D.

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLV)

DNA Sequencing by: Institute for Systems Biology

http://www.systemsbio.org

contact: amadan@systemsbiology.org

Anup Madan, Jeessica Fahy, Erin Helton, Mark Kettelman, Anuradha

Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

REMARK

COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAP Plate: 113 Row: e Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES
source
Location/Qualifiers
1. 657

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1430210"
/issue_type="Mammary gland"
/clone_lib="Soares_mammary_gland_NbMWG"
/lab_host="DH10B"
/note="Vector: pT73D-Pac"
BASE COUNT 159 a 173 c 190 g 135 t
ORIGIN

Query Match 2.0%; Score 26; DB 10; Length 657;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1292 CTCACGTAAAAAAGAAAAA 1317
628 CTCACGTAAAAAAGAAAAA 653

RESULT 32 660 bp mRNA linear STS 21-DEC-2001
LOCUS Snu171 Stickleback cDNA Gasterosteus aculeatus STS cDNA clone
DEFINITION 53-3/T7, sequence tagged site.
ACCESSION G72308
VERSION G72308.1 GI:17977168
KEYWORDS STS.
SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Gasterosteus aculeatus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.

1 (bases 1 to 660)

Buehler, C.L., Nereng, K.S., Ohgi, K.A., Cole, B.L., Colosimo, P.F.,

Buehler, C.A., Schluter, D. and Kingsley, D.M.

The genetic architecture of divergence between threespine

stickleback species

Nature 414 (6866), 901-905 (2001)

11780061

Synonyms: Cigreen

Contact: David M. Kingsley

Dept of Developmental Biology/ Howard Hughes Medical Institute

Stanford University School of Medicine

8300 Beckman Center, 279 Campus Drive, Stanford, CA 94305-5329, USA

Email: kingsley@cmmg.stanford.edu

Primer A: TATGTCCACGCTGCTCG

Primer B: AATGCTGTGTGATGACG

STS size: 120

PCR Profile:
Denaturation: 95 degrees C for 1:45
Annealing: 56 degrees C for 0:45
Polymerization: 72 degrees C for 0:45

PCR cycles: 1
Denaturation: 94 degrees C for 0:45
Annealing: 56 degrees C for 0:45
Polymerization: 72 degrees C for 0:45

PCR cycles: 4
Denaturation: 90 degrees C for 0:45
Annealing: 56 degrees C for 0:45
Polymerization: 72 degrees C for 0:45

PCR cycles: 30

Thermal cycler: MJ Research PTC-200 DNA Engine
Protocol:
Template: 5 ng
each 0.5 uM
DNTPs: each 62.5 uM
Tag polymerase: 0.25 units
Total volume: 10 uL

Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3
Location/Qualifiers
1. 660

/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas River, California"
/db_xref="taxon:69293"
/clone="53-3/T7"
/issue_type="Head and internal organs"
/clone_lib="Stickleback cDNA"
/dev_stage="adult"
/lab_host="XLI-Blue MRP"
/note="Vector: ZAP Express; V-type: Phage; Oligo(dT) primed cDNA unidirectionally cloned into the EcoRI/XhoI sites of ZAP Express."

STS
primer_bind 42. 61
primer_bind complement(142. 161)
BASE COUNT 209 a 144 c 100 g 193 t 14 others
ORIGIN

Query Match 2.0%; Score 26; DB 11; Length 660;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAGAAAAAAGCTGAG 1323
26 TAAAAAAGAAAAAAGCTGAG 1

RESULT 33 782 bp mRNA linear MAM 30-MAR-2003
LOCUS Sus scrofa partial mRNA for UDP-sugar diphosphatase (UGPP gene).
ACCESSION AF539380
VERSION AF539380.1 GI:29373949
KEYWORDS UDP-sugar diphosphatase; UGPP gene.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 Toshitiro, Y., Baroja-Fernandez, E., Yamamoto, R., Munoz, F.J.,

Akazawa, T., Hong, K.S. and Pozueta-Romero, J.

Cloning, expression and characterization of a mammalian Nudix

hydrolyase-like enzyme that cleaves the pyrophosphate bond of

UDP-glucose

Unpublished

2 (bases 1 to 782)

Pozueta-Romero, J.

Direct Submission

Submitted (30-JAN-2003) Pozueta-Romero, J., Universidad Publica de

Navarra / CSIC, Instituto de Agrobiotecnologia y Recurso, Carretera

de Mutilva s/n, 31192 Mutilva Beja (Navarra), SPAIN

Location/Qualifiers

1. 782

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

1. 782

/gene="UGPP"

gene

CDS

<1.671
/gene="UGPP"
/EC number="3.6.1.45"
/product="UDP-sugar diphosphatase"
/codon_start=3
/protein_id="CAD62374.1"
/db_xref="GI:29373950"
/translation="MERIEGAAGRCADSPYLRPLTLHYRQNGTOKSWDFVKTDSVT
LHMNASRSLVVKOPRPAVYAGAVRLFPGSLVADODRPRELPALPGSAGTVE
LCAGLVDPGLSLVEVCKEAMEECGYRLPSDLRVAYSKSGVGLTSSQTMFYAV
TDNRGPGGGLAESEGLIEVHLPLDGAQAFADDPVPKTLGVITGISFLSCVAV
LgPQ"

BASE COUNT 127 a 255 c 265 g 135 t

ORIGIN

Query Match 2.0%; Score 26; DB 4; Length 782;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
|||||
751 TAAAAAAAAAAAAAAAAAATCGAG 776

RESULT 34
A70165 786 bp DNA linear PAT 07-MAY-1999
LOCUS Sequence 26 from Patent WO9811228.
DEFINITION A70165
ACCESSION A70165
VERSION A70165.1 GI:4774580
KEYWORDS
SOURCE unidentified
ORGANISM unidentified

REFERENCE
1 (bases 1 to 786)
AUTHORS Seymour, G.B., Bird, C.R. and Medina-Suarez, R.D.
TITLE GENETIC CONTROL OF FRUIT RIPENING
JOURNAL Patent: WO 9811228-A 26 19-MAR-1998;
SEYMOUR GRAHAM BARON (GB)

FEATURES
Location/Qualifiers
source 1..786
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/clone="U-U125"

BASE COUNT 187 a 216 c 146 g 228 t 9 others

ORIGIN

Query Match 2.0%; Score 26; DB 6; Length 786;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
|||||
39 TAAAAAAAAAAAAAAAAAATCGAG 14

RESULT 35
AR254343 815 bp DNA linear PAT 20-DEC-2002
LOCUS Sequence 21 from patent US 6479733.
DEFINITION AR254343
ACCESSION AR254343
VERSION AR254343.1 GI:27303120
KEYWORDS
SOURCE Unknown.
ORGANISM Unidentified.

REFERENCE
1 (bases 1 to 815)
AUTHORS Rafalekt, J.A. and Famodu, O.O.
TITLE Sterol metabolism enzymes
JOURNAL Patent: US 6479733-A 21 12-NOV-2002;
FEATURES Location/Qualifiers
source 1..815.

BASE COUNT 206 a 173 c 200 g 236 t

ORIGIN

Query Match 2.0%; Score 26; DB 6; Length 815;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
|||||
628 TAAAAAAAAAAAAAAAAAATCGAG 653

RESULT 36
BD1422213 1038 bp mRNA linear INV 18-FEB-2002
LOCUS BD1422213
DEFINITION Babesia divergens mRNA for ran binding-like protein 1.
ACCESSION AJ422213
VERSION AJ422213.1 GI:18857688
KEYWORDS ran binding-like protein 1.
SOURCE Babesia divergens
ORGANISM Babesia divergens
Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae;
Babesia.

REFERENCE
1
AUTHORS Delbecq, S.
TITLE Babesia divergens : cloning of a Ran Binding Protein 1 homologue in
an intrerythrocytic protozoa
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1038)
TITLE Delbecq, S.
JOURNAL Direct Submission
Submitted (18-DEC-2001) Delbecq S., Laboratoire de Biologie
Cellulaire et Mo, Faculte de Pharmacie, 15, av. C. Flahault B.P.
14491, 34093 Montpellier Cedex 5, FRANCE

FEATURES
Location/Qualifiers
source 1..1038
/organism="Babesia divergens"
/mol_type="mRNA"
/isolate="Rouen 1987"
/db_xref="taxon:32595"
/country="France"
129..731
/function="nucleocytoplasmic transport"
/codon_start=1
/product="ran binding-like protein 1"
/protein_id="CAD19562.1"
/db_xref="GI:18857689"
/translation="MTDLAKEDPNVPSDVTTEPTPKDAQCEPKTEAAQPEDAQOE
EEEVGDMTRKVEVKEIKVVTGEDEEIFMOQREKLYRFGDTGDQDQVWKGKGLGE
AKIKKTKTGKTRILROEKTLKVVANHYVLEALCKLPKNGSKINWWTANNYYD
DETKVQMLKRGQIEQAKLFEKEFEAAKINKATFDKE"

BASE COUNT 335 a 229 c 284 g 190 t

ORIGIN

Query Match 2.0%; Score 26; DB 3; Length 1038;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
|||||
885 TAAAAAAAAAAAAAAAAAATCGAG 910

RESULT 37
AR243822 1126 bp DNA linear PAT 20-DEC-2002
LOCUS AR243822
DEFINITION Sequence 60 from patent US 6476195.
ACCESSION AR243822
VERSION AR243822.1 GI:27291315
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1126)
AUTHORS Komatsoulis, G., Rosen, C.A., Ruben, S.M., Duan, R.D., Moore, P.A., Shi, Y., LaFleur, D.W., Wei, Y.-F., Ni, J., Florence, K.A., Young, P., Brewer, L.A., Soppet, D.R., Endress, G.A., Ebner, R., Olsen, H. and Mucenaiski, M.
TITLE Secreted protein HNFGE20
JOURNAL Patent: US 6476195-A 60 05-NOV-2002;
FEATURES location/Qualifiers
source 1..1126
/organism="unknown"
BASE COUNT 266 a 306 c 292 g 257 t 5 others
ORIGIN
Query Match 2.0%; Score 26; DB 6; Length 1126;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1101 TAAAAAAAAAAAAAAAACTCGAG 1126
RESULT 38
AR307577 1141 bp DNA linear PAT 12-JUN-2003
LOCUS AR307577
DEFINITION Sequence 7 from patent US 6551807.
ACCESSION AR307577
VERSION AR307577.1 GI:31698167
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1141)
AUTHORS Cunningham, F.X.
TITLE Carotenoid ketolase genes and gene products, production of the ketocarotenoids and methods of modifying carotenoids using the genes
JOURNAL Patent: US 6551807-A 7 22-APR-2003;
FEATURES location/Qualifiers
source 1..1141
/organism="unknown"
BASE COUNT 330 a 234 c 247 g 330 t
ORIGIN
Query Match 2.0%; Score 26; DB 6; Length 1141;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1116 TAAAAAAAAAAAAAAAACTCGAG 1141
RESULT 39
BD140651 1149 bp DNA linear PAT 18-SEP-2002
LOCUS BD140651
DEFINITION 36 human secreted proteins.
ACCESSION BD140651.1 GI:23235596
VERSION BD140651.1 GI:23235596
KEYWORDS JP 2002504361-A/37.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1149)
AUTHORS Olsen, H.S., Florence, K., Brewer, L.A., Ebner, R., Ruben, S.M., Rosen, C.A., and Duan, R.D.
TITLE 36 human secreted proteins
JOURNAL Patent: JP 2002504361-A 37 12-FEB-2002;
COMMENT OS Homo sapiens (human)
PN JP 2002504361-A/37

PD 12-FEB-2002
PF 24-FEB-1999 JP 2000533445
PR 26-FEB-1998 US 60/076051, 26-FEB-1998 US 60/076052 PR
26-FEB-1998 US 60/076053, 26-FEB-1998 US 60/076054 PR
26-FEB-1998 US 60/076057
PI HENRIK S OLSEN, KIMBERLY FLORENCE, LAURIE A BREWER, REINHARD P
EBNER, STEVEN M RUBEN, CRAIG A ROSEN, ROXANNE D DUAN
PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P7/00, A61P11/06,
PC A61P15/00,
PC A61P19/02, A61P25/00, A61P29/00, A61P37/06, C07H21/00, C07K14/47,
PC C07K16/18,
C12N1/15, C12N1/21, C12N5/10, G01N33/68, C12N15/00, A61K37/02, C12N5/
00
CC 36 human secreted proteins
FH Key location/Qualifiers
FT source 1..1149
FT location/Qualifiers
/organism="Homo sapiens (human)".
FEATURES source
location/Qualifiers
1..1149
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 374 a 176 c 195 g 404 t
ORIGIN
Query Match 2.0%; Score 26; DB 6; Length 1149;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1124 TAAAAAAAAAAAAAAAACTCGAG 1149
RESULT 40
GGA421070 1151 bp mRNA linear VRT 30-NOV-2001
LOCUS GGA421070
DEFINITION Gallus gallus mRNA for bHLH transcription factor (CMES02 gene).
ACCESSION A421070
VERSION A421070.1 GI:17221420
KEYWORDS bHLH transcription factor; CMES02 gene.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianidae; Gallus.
REFERENCE 1
1 Buchberger, A., Bonnelick, S., Klein, C. and Arnold, H.H.
Dynamic expression of chicken cmes02 in segmental plate and somites
Dev. Dyn.
2 (bases 1 to 1151)
AUTHORS Buchberger, A.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-2001) Buchberger A., Cell and Molecular Biology,
TU Braunschweig, Spielmannstrasse 7, Braunschweig, 38106, GERMANY
FEATURES location/Qualifiers
source 1..1151
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/tissue_type="whole embryo"
/dev_stage="HH stage 8"
1..1151
/gene="CMES02"
1..972
/gene="CMES02"
/codon_start=1
/product="bHLH transcription factor"
/protein_id="CAD1282.1"
/db_xref="GI:17221421"
/translation="MAGSPLLHGLQAPTTASLTLPPELCPGSGAGLRRCGPAGRR

GSQVAGPQASEREKLRMRRLAQMHRLHYLPALAPACONTXIKETLRATRYI
AHLISALGLSEBVLLRGAPAPRHCPCCGGLCCOPPRRLPLAPQASPGVLGM
GTPLELGLAGTGTWGLSGSPVLETSPGLATGMLSPSPPTAGASPEL
HGTGSMQSPSPPLCSIRAPETRAETACTATGTATGTCMCEAAASLHPGNMAK
AMGSPSLPTGSGSIGSPDLDMFISDAGCTPVLPGCSFARRRGNWS"

BASE COUNT 200 a 416 c 373 g 162 t

ORIGIN

Query Match 2.0%; Score 26; DB 5; Length 1151;
Best Local Similarity 100.0%; Pred. No. 0.082; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1069 TAAAAAAAAAAAAAAAACTCGAG 1094

RESULT 41
BD136379 1263 bp DNA linear PAT 18-SEP-2002
LOCUS 95 human secretory proteins.
ACCESSION BD136379
VERSION BD136379.1 GI:23231324
KEYWORDS JP 2002506627-A/66.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1263)
Ruden,S.M., Ni,J., Rosen,C.A., Yu,G.L., Young,P.E., Fen,P.,
Soppet,D.R., Wei,Y.F., Endress,G.A., Duan,R.D., Kyaw,H., Ebner,R.,
Lafleur,D.W., Olsen,H.S., Shi,Y. and Moore,P.A.
95 human secretory proteins
Patent: JP 2002506627-A 66 05-MAR-2002;
HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
PN JP 2002506627-A/66
PD 05-MAR-2002
PR 18-MAR-1999 JP 2000536733
PR 19-MAR-1998 US 60/078566,19-MAR-1998 US 60/078574 PR
19-MAR-1998 US 60/078573,19-MAR-1998 US 60/078574 PR
19-MAR-1998 US 60/078579,19-MAR-1998 US 60/078578 PR
19-MAR-1998 US 60/078581,19-MAR-1998 US 60/078577 PR
19-MAR-1998 US 60/078563,01-APR-1998 US 60/080314 PR
01-APR-1998 US 60/080312,01-APR-1998 US 60/080313 PI STEVEN
M RUBEN,JIAN NI,CRAIG A ROSEN,GUO
LIANG YU,PAUL E YOUNG,
PI PING FENG,
PI DANIEL R SOPPET,YING FEI WEI,GREGORY A ENDRESS,ROXANNE D DUAN,
PI HLA KYAM,
PI REINHARD EBNER,DAVID W LAFLEUR,HENRIK S OLSEN,YANGGU SHI,PAUL
PI A MOORE
PC C12N15/09,A61K38/00,A61K48/00,A61P43/00,C07K14/47,C07K16/18,
PC C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12Q1/68,G01N33/53,G01N33/PC
566,
CC n equals a,t,g, or c
CC n equals a,t,g, or c
FH Key Location/Qualifiers
FT source 1..1263
FT Location/Qualifiers
FT /organism="Homo sapiens (human)"

FEATURES
source 1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 303 a 272 c 290 g 397 t 1 others

ORIGIN

Query Match 2.0%; Score 26; DB 6; Length 1263;
Best Local Similarity 100.0%; Pred. No. 0.081; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1100 TAAAAAAAAAAAAAAAACTCGAG 1125

RESULT 42
AR225426 1331 bp DNA linear PAT 20-DEC-2002
LOCUS Sequence 27 from patent US 6444425.
ACCESSION AR225426
VERSION AR225426.1 GI:27263372
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1331)
Reed,S.G., Lodes,M.J., Mohamath,R. and Secrist,H.
Compounds for therapy and diagnosis of lung cancer and methods for
their use
Patent: US 6444425-A 27 03-SEP-2002;
Location/Qualifiers
FEATURES source 1..1331
/organism="unknown"
BASE COUNT 409 a 270 c 304 g 348 t

ORIGIN

Query Match 2.0%; Score 26; DB 6; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0.081; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1306 TAAAAAAAAAAAAAAAACTCGAG 1331

RESULT 43
AX321496 1331 bp DNA linear PAT 15-DEC-2001
LOCUS Sequence 27 from Patent WO0172295.
ACCESSION AX321496
VERSION AX321496.1 GI:17905560
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
Indrias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
Manton,J. and Kaios,M.D.
Compositions and methods for the therapy and diagnosis of lung
cancer
Patent: WO 0172295-A 27 04-OCT-2001;
CORIXA CORPORATION (US)
Location/Qualifiers
FEATURES source 1..1331
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 409 a 270 c 304 g 348 t

ORIGIN

Query Match 2.0%; Score 26; DB 6; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0.081; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1306 TAAAAAAAAAAAAAAAACTCGAG 1331

RESULT 44

AB058679/c
LOCUS AB058679 1351 bp mRNA linear PLN 29-MAR-2001
DEFINITION Pyrus communis PY-PIPI-1 mRNA for plasma membrane intrinsic protein 1-1, complete cds.
ACCESSION AB058679
VERSION AB058679.1 GI:13486939
KEYWORDS Pyrus communis (pear)
SOURCE Pyrus communis
ORGANISM Pyrus communis
REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Rosales; Rosaceae; Maloideae; Pyrus.
1 (sites)
AUTHORS Kobae, Y., Shiratake, K., Goto, H. and Yamaki, S.
TITLE Molecular cloning of cDNAs encoding plasma membrane aquaporins of pear fruit
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 1351)
AUTHORS Kobae, Y., Shiratake, K., Goto, H. and Yamaki, S.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2001) Katsuhito Shiratake, Nagoya University, Graduate School of Bio-agricultural Sciences, Chikusa, Nagoya, Aichi 464-8601, Japan (E-mail:shiratakegr.nagoya-u.ac.jp, Tel:81527894027, Fax:81527894025)
FEATURES
source location/Qualifiers
1. .1351
/organism="Pyrus communis"
/mol_type="mRNA"
/db_xref="taxon:23211"
1. .1351
/gene="PY-PIPI-1"
225. .1094
/gene="PY-PIPI-1"
/codon_start=1
/product="plasma membrane intrinsic protein 1-1"
/protein_id="BAB40142.1"
/db_xref="GI:13486940"
/translation="MEGKEDVRLGANKFSEKROPIGSAOTDEGDKYKEPPAPLPE
PGLTMSFYRAGIAEFLATFLYITLTMGVKSKCTTGIGIIMARCGTIF
ALVYSTGIGSGHINPAVTGFLPRLKSLTRAVVMTOTLAINAAVVKPKES
TFEMLGGAUSVAHGTYKGGGLGAEITGTFLVTVTSATDKRSADSVPIALPI
IGFAVFLHATLPIVGTGTGINPARSLGAAITYNKOHAWDHWFWGPFGLAALALY
HVVVIRAIPIKSKK"
BASE COUNT 320 a 335 c 338 g 358 t
ORIGIN
Query Match 2.0%; Score 26; DB 8; Length 1351;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAATCTGAG 1323
Db 31 TAAAAAATCTGAG 6
RESULT 45
LOCUS AR071757 1352 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 44 from patent US 5912140.
ACCESSION AR071757
VERSION AR071757.1 GI:7222645
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1352)
AUTHORS Whoriskey, S.K., Quinn, C.L., Tao, N., Politis-Viark, K.I. and Schimmel, P.R.
TITLE Recombinant pneumocystis carinii aminocycl tRNA synthetase genes, tester strains and assays
JOURNAL Patent: US 5912140-A 44 15-JUN-1999;
FEATURES
source location/Qualifiers
1. .1352

BASE COUNT 519 a 185 c 217 g 431 t
ORIGIN
Query Match 2.0%; Score 26; DB 6; Length 1352;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAATCTGAG 1323
Db 1327 TAAAAAATCTGAG 1352
RESULT 46
LOCUS BD136397 1433 bp DNA linear PAT 18-SEP-2002
DEFINITION 95 human secretory proteins.
ACCESSION BD136397
VERSION BD136397.1 GI:23211342
KEYWORDS JP 2002506627-A/84.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1433)
AUTHORS Ruben, S.M., Ni, J., Rosen, C.A., Yu, G.L., Young, P.E., Fen, P., Soppet, D.R., Wei, Y.F., Endress, G.A., Duan, R.D., Kyaw, H., Ebner, R., Lafleur, D.W., Olsen, H.S., Shi, Y. and Moore, P.A.
TITLE 95 human secretory proteins
JOURNAL Patent: JP 2002506627-A 84 05-MAR-2002;
COMMENT HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002506627-A/84
PD 05-MAR-2002
PF 18-MAR-1998 JP 200053573
PR 19-MAR-1998 US 60/078566, 19-MAR-1998 US 60/078574 PR
19-MAR-1998 US 60/078573, 19-MAR-1998 US 60/078578 PR
19-MAR-1998 US 60/078581, 19-MAR-1998 US 60/078578 PR
19-MAR-1998 US 60/078563, 01-APR-1998 US 60/080314 PR
01-APR-1998 US 60/080312, 01-APR-1998 US 60/080313 PR STEVEN
M RUBEN, JIAN NI, CRAIG A ROSEN, GUO
LIANG YU, PAUL E YOUNG,
PING FENG,
DANIEL R SOPPET, YING FEI WEI, GREGORY A ENDRESS, ROXANNE D DUAN,
HIA KYAW,
PI REINHARD EBNER, DAVID W LAFLEUR, HENRIK S OLSEN, YANGU SHI, PAUL
PI A MOORE
PC C12N15/09, A61K38/00, A61K48/00, A61P43/00, C07K14/47, C07K16/18,
PC C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N33/53, G01N33/ PC
566,
PC C12N15/00, A61K37/02, C12N5/00
CC 95 human secretory proteins
FH key location/Qualifiers
FT source 1. .1433
/organism="Homo sapiens (human)".
FEATURES
source location/Qualifiers
1. .1433
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 291 a 433 c 417 g 290 t 2 others
ORIGIN
Query Match 2.0%; Score 26; DB 6; Length 1433;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAATCTGAG 1323
Db 1392 TAAAAAATCTGAG 1417

RESULT 47
PCUS7795 1433 bp mRNA linear PLN 29-DEC-1998
LOCUS Pneumocystis carinii f. sp. ratii S-adenosylhomocysteine hydrolase
DEFINITION (SAHH) mRNA, partial cds.
ACCESSION U57795
VERSION GI:1401141
KEYWORDS
SOURCE Pneumocystis carinii f. sp. ratii
ORGANISM Pneumocystis carinii f. sp. ratii
REFERENCE 1 (bases 1 to 1433)
AUTHORS Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
Pneumocystidaceae; Pneumocystis.
Lasbury M.E., Brady S., McLaughlin G., Bartlett M.S., Smith J.W.
TITLE Pneumocystis carinii ratius mRNA sequence for
S-adenosylhomocysteine hydrolase
JOURNAL 2 (bases 1 to 1433)
AUTHORS Unpublished
Lasbury M.E., Brady S., McLaughlin G., Bartlett M.S., Smith J.W.
and Lee C.H.
REFERENCE Direct Submission
AUTHORS Submitted (08-MAY-1996) Pathology, Indiana University, 1120 South
Drive, Fessler Hall, Rm. 419, Indianapolis, IN 46202, USA
LOCATION/Qualifiers
FEATURES
source
1. 1433
/organism="Pneumocystis carinii f. sp. ratii"
/mol_type="mRNA"
/db_xref="taxon:38082"
/note="forma specialis=ratii"
1. 1433
/gene="SAHH"
1. 1323
/gene="SAHH"
/codon_start=1
/product="S-adenosylhomocysteine hydrolase"
/protein_id="AAC98514.1"
/db_xref="GI:1401142"
/translation="SCRFTLAEWGRKEIELAENEMGLISCRNMFKPLRVNACGL
HMTIQVAIVLELVELGAETVSSCNIPSTODHAAAIASGISVAMKETEELW
CIETQLTSPFDGHLNMLDDGDVTLGVNKKYPDVLAOKGISEBTGGVHVFYML
KEGKLKPAIVNVDVTKSKFDNLGVLNKKYISNLIKRCRESLIDGIKRATDIM
IAKVAIVAGFDVGKCAKALRGMGARVITTEIDPVALOASMEGOVAVMEVAD
ADIFVATGCKDIDICERHPEAMKNDALICNIGHFDVIDAMLIKCSISNKKPOVD
RYLGNRNITLLAEGRLVNLGCATGSHSSVMWSTNOVALAIALMTAOGQYPLGV
HPLPKLDEVARHLHSLKGLKLSLTPESQAVLDIPDGYSKSHRY"

BASE COUNT 487 a 206 c 289 g 451 t
ORIGIN

Query Match 2.0%; Score 26; DB 8; Length 1433;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATACTCGAG 1323
DB 1407 TAAAAAATACTCGAG 1432

RESULT 48
S47891 1467 bp mRNA linear VRT 08-MAY-1993
LOCUS XSTR2activin receptor [Xenopus laevis, mRNA, 1467 nt].
DEFINITION S47891
ACCESSION S47891
VERSION GI:259178
KEYWORDS
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 1467)
LOCATION/Qualifiers

AUTHORS Nishimatsu, S., Iwao, M., Nagai, T., Oda, S., Suzuki, A., Asashima, M.,
Murakami, K. and Ueno, N.
TITLE A carboxyl-terminal truncated version of the activin receptor
JOURNAL mediates activin signals in early Xenopus embryos
MEDLINE FEBS Lett. 312 (2-3), 169-173 (1992)
PUBMED 93050187
REMARK Genbank staff at the National Library of Medicine created this
entry [NCBI gi246117250] from the original journal article.
This sequence comes from Fig. 1.
LOCATION/Qualifiers
FEATURES
source
1. 1467
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
1. 1467
/gene="XSTR2"
/note="activin receptor"
325. 1422
/gene="XSTR2"
/note="This sequence comes from Fig. 1"
/product="activin receptor"
/protein_id="AA024030.1"
/db_xref="GI:259178"
/translation="MGAAYPLTALATLAFRAGSGHDEVEYRECIYANWMELEKTNQ
SGVSECEGEKDRHLHCYASMRNNSGFIELVKGCMWDPNDCYDRQICANKEPQVFF
CCCEBNYCNKFTPLBEVETFDPKPQPMPSVNIILYSLPIAGLSMVLILAMPMYRH
RKRPYGVNDVNDPGRSPSPVAVKLPDOLLITKARGRCVWKARLLAEVAVKIP
VDKDSWQCEKEIFNTPGKHENLSEFIAEGRGSLNELMLITRPHDKGSTIDYK
GNLVWNLCHITETMARGLSYLHEDVPRCKEGEGRPAIAHDFKSNVLRINDLTAL
LDFGLAVFEFEPKPGDTHGVICAAINL"

BASE COUNT 381 a 354 c 390 g 342 t
ORIGIN

Query Match 2.0%; Score 26; DB 5; Length 1467;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATACTCGAG 1323
DB 1442 TAAAAAATACTCGAG 1467

RESULT 49
AF228334 1550 bp mRNA linear VRT 09-AUG-2000
LOCUS Danto rerio dHAND mRNA, complete cds.
DEFINITION AF228334
ACCESSION AF228334
VERSION GI:7677350
KEYWORDS
SOURCE Danto rerio (zebrafish)
ORGANISM Danto rerio
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 1550)
Yost, H.J., Lohr, J., Lee, K.H., Ticho, B.S., Breitbart, R.E., Hill, S.,
Yost, H.J., and Srivastava, D.
TITLE Conservation of sequence and expression of Xenopus and zebrafish
dHAND during cardiac, branchial arch and lateral mesoderm
development
JOURNAL Mech. Dev. 95 (1-2), 231-237 (2000)
MEDLINE 20368175
PUBMED 10906469
REFERENCE 2 (bases 1 to 1550)
AUTHORS Angelo, S., Lohr, J., Lee, K.H., Ticho, B.S., Breitbart, R.E., Hill, S.,
Yost, H.J., and Srivastava, D.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-2000) Pediatrics/Molecular Biology, University of
Texas Southwestern Medical Center, 6000 Harry Hines Blvd, Dallas,
TX 75390, USA
LOCATION/Qualifiers

source 1..1550
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 252..878
 /codon_start=1
 /product="dHAND"
 /protein_id="AA067130.1"
 /db_xref="GI:7677351"
 /translation="MSLVGGPRPHVPHVHNDGYSFAAASRCHPEPPYFHWLISHP
 EMSPPDYTMAPSISPEYSTGAPGLDHSHYGVGAGAVGMPRTYKRRTANKRERR
 TQSLNSAFELRECTPNVPADTKLSIKTLRLATSYLAIMDLIDKDEQNGEYFAFRA
 EFKTDKERERRKKNMDVLKSSGSSNDKKTGRTGMPQHWMALELKQ"

BASE COUNT 457 a 322 c 331 g 440 t
 ORIGIN

Query Match 2.0%; Score 26; DB 5; Length 1550;
 Best Local Similarity 100.0%; Pred. No. 0.08;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATCTCGAG 1323
 Db 1410 TAAAAAATCTCGAG 1435

RESULT 50
 AX074228 1579 bp DNA linear PAT 06-FEB-2001
 LOCUS
 DEFINITION Sequence 11 from Patent WO0104147.
 AX074228
 ACCESSION
 VERSION AX074228.1 GI:12710426
 KEYWORDS
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE
 AUTHORS Cahoon, R.E.
 TITLE Plant inositol polyphosphate phosphatase homologs
 JOURNAL Patent: WO 0104147-A 11 18-JAN-2001;
 E.I. DU PONT DE NEMOURS AND COMPANY (US)
 FEATURES
 source 1..1579
 location/Qualifiers
 /organism="Triticum aestivum"
 /mol_type="genomic DNA"
 /db_xref="taxon:4565"

BASE COUNT 450 a 356 c 333 g 440 t
 ORIGIN

Query Match 2.0%; Score 26; DB 6; Length 1579;
 Best Local Similarity 100.0%; Pred. No. 0.08;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATCTCGAG 1323
 Db 1524 TAAAAAATCTCGAG 1549

Search completed: January 28, 2004, 17:25:12
 Job time : 4879 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 28, 2004, 18:19:00 ; Search time 44 Seconds
(without alignments)
185.591 Million cell updates/sec

Title: US-09-881-556A-4
Perfect score: 193
Sequence: 1 HEIFFLPLRISIFSDSND.....KAPIKENGLNNGSKYK 193

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 5

Total number of hits satisfying chosen parameters: 4387

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A COMB .pep.*
2: /cgn2_6/prodata/2/1aa/5B COMB .pep.*
3: /cgn2_6/prodata/2/1aa/6A COMB .pep.*
4: /cgn2_6/prodata/2/1aa/6B COMB .pep.*
5: /cgn2_6/prodata/2/1aa/PCTUS COMB .pep.*
6: /cgn2_6/prodata/2/1aa/backfile1 .pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	4.1	520	4 US-09-328-352-6592	Sequence 6592, Ap
2	7	3.6	72	4 US-09-663-600A-93	Sequence 93, Appl
3	7	3.6	72	4 US-09-663-600A-187	Sequence 187, Appl
4	7	3.6	100	4 US-09-152-060-61	Sequence 61, Appl
5	7	3.6	106	4 US-09-663-600A-95	Sequence 95, Appl
6	7	3.6	106	4 US-09-663-600A-189	Sequence 189, Appl
7	7	3.6	212	4 US-09-702-705-1820	Sequence 1820, Ap
8	7	3.6	212	4 US-09-726-457-1820	Sequence 1820, Ap
9	7	3.6	223	4 US-09-230-196-17	Sequence 17, Appl
10	7	3.6	223	4 US-09-230-196-24	Sequence 24, Appl
11	7	3.6	282	3 US-08-326-728A-55	Sequence 56, Appl
12	7	3.6	432	4 US-09-328-352-6119	Sequence 6119, Ap
13	7	3.6	504	1 US-08-441-139-18	Sequence 18, Appl
14	7	3.6	3169	4 US-09-453-702B-257	Sequence 257, Appl
15	6	3.1	8	4 US-08-817-832B-26	Sequence 26, Appl
16	6	3.1	26	2 US-08-620-151-52	Sequence 52, Appl
17	6	3.1	32	1 US-08-118-370-186	Sequence 186, Appl
18	6	3.1	32	1 US-08-118-370-188	Sequence 188, Appl
19	6	3.1	32	5 PCT-US93-08528-186	Sequence 186, Appl
20	6	3.1	32	5 PCT-US93-08528-188	Sequence 188, Appl
21	6	3.1	37	4 US-08-331-625A-20	Sequence 20, Appl
22	6	3.1	37	4 US-09-494-151-20	Sequence 20, Appl
23	6	3.1	38	5 PCT-US93-08528-187	Sequence 187, Appl
24	6	3.1	38	5 PCT-US93-08528-187	Sequence 187, Appl
25	6	3.1	44	4 US-09-266-764-24	Sequence 24, Appl
26	6	3.1	46	4 US-08-861-476C-5	Sequence 5, Appl
27	6	3.1	67	4 US-09-252-991A-22890	Sequence 22890, A

28	6	3.1	68	4 US-09-134-001C-4520	Sequence 4520, Ap
29	6	3.1	74	4 US-09-205-258-1151	Sequence 1151, Ap
30	6	3.1	78	3 US-09-073-297-18	Sequence 18, Appl
31	6	3.1	81	4 US-09-205-258-1038	Sequence 1038, Ap
32	6	3.1	86	4 US-09-375-140-9	Sequence 9, Appl
33	6	3.1	90	3 US-08-665-259-2	Sequence 2, Appl
34	6	3.1	90	3 US-08-762-500-2	Sequence 2, Appl
35	6	3.1	98	4 US-08-936-165A-393	Sequence 393, Appl
36	6	3.1	126	3 US-08-331-625A-9	Sequence 9, Appl
37	6	3.1	126	4 US-09-494-151-9	Sequence 9, Appl
38	6	3.1	127	3 US-08-905-223-310	Sequence 310, Appl
39	6	3.1	128	4 US-08-858-207A-314	Sequence 314, Appl
40	6	3.1	141	4 US-09-252-991A-3648	Sequence 3648, A
41	6	3.1	145	2 US-08-788-584-1	Sequence 1, Appl
42	6	3.1	149	2 US-08-788-584-3	Sequence 3, Appl
43	6	3.1	155	4 US-09-198-452A-507	Sequence 507, Appl
44	6	3.1	157	4 US-09-162-021B-10	Sequence 10, Appl
45	6	3.1	164	2 US-08-484-993B-45	Sequence 45, Appl
46	6	3.1	164	2 US-08-484-158B-45	Sequence 45, Appl
47	6	3.1	164	2 US-08-484-596A-45	Sequence 45, Appl
48	6	3.1	164	2 US-08-480-150A-45	Sequence 45, Appl
49	6	3.1	164	3 US-08-458-731-45	Sequence 45, Appl
50	6	3.1	164	3 US-08-149-222A-45	Sequence 45, Appl
51	6	3.1	168	4 US-09-328-352-5484	Sequence 5484, Ap
52	6	3.1	192	4 US-09-134-001C-5386	Sequence 5386, Ap
53	6	3.1	196	4 US-09-328-352-6310	Sequence 6310, Ap
54	6	3.1	197	3 US-09-112-248-2	Sequence 2, Appl
55	6	3.1	197	4 US-09-162-021B-8	Sequence 8, Appl
56	6	3.1	197	4 US-09-687-477-8	Sequence 8, Appl
57	6	3.1	197	4 US-09-687-477-10	Sequence 10, Appl
58	6	3.1	197	4 US-09-687-477-12	Sequence 12, Appl
59	6	3.1	197	4 US-09-687-477-14	Sequence 14, Appl
60	6	3.1	197	4 US-09-687-477-16	Sequence 16, Appl
61	6	3.1	197	4 US-09-687-477-18	Sequence 18, Appl
62	6	3.1	197	4 US-09-687-476-10	Sequence 10, Appl
63	6	3.1	197	4 US-09-687-476-12	Sequence 12, Appl
64	6	3.1	197	4 US-09-687-476-14	Sequence 14, Appl
65	6	3.1	197	4 US-09-687-476-16	Sequence 16, Appl
66	6	3.1	197	4 US-09-687-372-8	Sequence 8, Appl
67	6	3.1	197	4 US-09-687-372-10	Sequence 10, Appl
68	6	3.1	197	4 US-09-687-372-12	Sequence 12, Appl
69	6	3.1	197	4 US-09-687-372-14	Sequence 14, Appl
70	6	3.1	197	4 US-09-687-372-16	Sequence 16, Appl
71	6	3.1	197	4 US-09-975-553-8	Sequence 8, Appl
72	6	3.1	197	4 US-09-975-553-10	Sequence 10, Appl
73	6	3.1	197	4 US-09-975-553-12	Sequence 12, Appl
74	6	3.1	197	4 US-09-975-553-14	Sequence 14, Appl
75	6	3.1	197	4 US-09-975-553-16	Sequence 16, Appl
76	6	3.1	198	2 US-08-943-915-33	Sequence 33, Appl
77	6	3.1	198	4 US-09-162-021B-6	Sequence 6, Appl
78	6	3.1	199	4 US-09-687-477-2	Sequence 2, Appl
79	6	3.1	199	4 US-09-687-477-4	Sequence 4, Appl
80	6	3.1	199	4 US-09-687-477-6	Sequence 6, Appl
81	6	3.1	199	4 US-09-687-477-8	Sequence 8, Appl
82	6	3.1	199	4 US-09-687-476-2	Sequence 2, Appl
83	6	3.1	199	4 US-09-687-476-4	Sequence 4, Appl
84	6	3.1	199	4 US-09-687-476-6	Sequence 6, Appl
85	6	3.1	199	4 US-09-687-476-8	Sequence 8, Appl
86	6	3.1	199	4 US-09-687-372-2	Sequence 2, Appl
87	6	3.1	199	4 US-09-687-372-4	Sequence 4, Appl
88	6	3.1	199	4 US-09-687-372-6	Sequence 6, Appl
89	6	3.1	199	4 US-09-687-372-8	Sequence 8, Appl
90	6	3.1	199	4 US-09-975-553-2	Sequence 2, Appl
91	6	3.1	199	4 US-09-975-553-4	Sequence 4, Appl
92	6	3.1	199	4 US-09-975-553-6	Sequence 6, Appl
93	6	3.1	199	4 US-09-975-553-8	Sequence 8, Appl
94	6	3.1	204	4 US-09-107-532A-5769	Sequence 5769, Ap
95	6	3.1	207	2 US-08-943-915-2	Sequence 2, Appl
96	6	3.1	207	2 US-08-943-915-5	Sequence 5, Appl
97	6	3.1	217	3 US-08-341-018-50	Sequence 50, Appl
98	6	3.1	217	4 US-08-467-602-413	Sequence 413, Appl
99	6	3.1	218	1 US-08-470-837-26	Sequence 26, Appl
100	6	3.1	218	4 US-08-868-452-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1

US-09-328-352-6592
; Sequence 6592, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Berton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6592
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6592

Query Match 4.1%; Score 8; DB 4; Length 520;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 TINKINKS 147
DB 3 TINKINKS 10

RESULT 2

US-09-663-600A-93
; Sequence 93, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Dugclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 93
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -32...-1
US-09-663-600A-93

Query Match 3.6%; Score 7; DB 4; Length 72;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 NKISLES 150
DB 59 NKISLES 65

RESULT 3

US-09-663-600A-187
; Sequence 187, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Dugclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 187
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -32...-1
US-09-663-600A-187

Query Match 3.6%; Score 7; DB 4; Length 72;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 NKISLES 150
DB 59 NKISLES 65

RESULT 4

US-09-152-060-61
; Sequence 61, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30

EARLIER APPLICATION NUMBER: 60/048,189
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/057,765
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/068,368
EARLIER FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 61
LENGTH: 100
TYPE: PRT
ORGANISM: Homo sapiens
US-09-152-060-61

Query Match 3.6%; Score 7; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 NKISLES 150
DB 53 NKISLES 59

RESULT 5
US-09-663-600A-95

Sequence 95, Application US/09663600A
Patent No. 6573068
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclercq, Aymeric
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
FILE REFERENCE: 31.US3.CIP
CURRENT APPLICATION NUMBER: US/09/663,600A
CURRENT FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/191,997
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: 60/066,677
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/069,957
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/074,121
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/081,563
PRIOR FILING DATE: 1998-04-13
PRIOR APPLICATION NUMBER: 60/096,116
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/099,273
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 229
SOFTWARE: Patent.pm
SEQ ID NO 95
LENGTH: 106
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: SIGNAL
LOCATION: -32...-1
US-09-663-600A-95

Query Match 3.6%; Score 7; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 NKISLES 150
DB 59 NKISLES 65

RESULT 6
US-09-663-600A-189

Sequence 189, Application US/09663600A
Patent No. 6573068
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclercq, Aymeric
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
FILE REFERENCE: 31.US3.CIP
CURRENT APPLICATION NUMBER: US/09/663,600A
CURRENT FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/191,997
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: 60/066,677
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/069,957
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/074,121
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/081,563
PRIOR FILING DATE: 1998-04-13
PRIOR APPLICATION NUMBER: 60/096,116
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/099,273
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 229
SOFTWARE: Patent.pm
SEQ ID NO 189
LENGTH: 106
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: SIGNAL
LOCATION: -32...-1
US-09-663-600A-189

Query Match 3.6%; Score 7; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 NKISLES 150
DB 59 NKISLES 65

RESULT 7

US-09-702-705-1820
Sequence 1820, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Ranger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1820
LENGTH: 212
TYPE: PRT
ORGANISM: Homo sapiens
US-09-702-705-1820

Query Match 3.6%; Score 7; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 41;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 84 CFEKNEA 90
Db 121 CFEKNEA 127

RESULT 8

US-09-736-457-1820
Sequence 1820, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaifanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fangert, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Pan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1820
LENGTH: 212
TYPE: PRT
ORGANISM: Homo sapiens
US-09-736-457-1820

Query Match 3.6%; Score 7; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 CFEKNEA 90
Db 121 CFEKNEA 127

RESULT 9

US-09-230-196-17
Sequence 17, Application US/09230196
Patent No. 6307035
GENERAL INFORMATION:
APPLICANT: Rauscher III, Frank J.
APPLICANT: Jensen, David E.
TITLE OF INVENTION: BRCA1 Associated Protein (BAP-1) and
TITLE OF INVENTION: Uses Therefor
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/230,196
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/022,997
FILING DATE: 02-AUG-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/038,109
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST68BUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-230-196-17

Query Match 3.6%; Score 7; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 CFEKNEA 90
Db 132 CFEKNEA 138

RESULT 10
US-09-230-196-24
Sequence 24, Application US/09230196
Patent No. 6307035
GENERAL INFORMATION:
APPLICANT: Rauscher III, Frank J.
APPLICANT: Jensen, David E.
TITLE OF INVENTION: BRCA1 Associated Protein (BAP-1) and
TITLE OF INVENTION: Uses Therefor
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/230,196
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/022,997
FILING DATE: 02-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/038,109
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST68BUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown

MOLECULE TYPE: protein
US-09-230-196-24

Query Match 3.6%; Score 7; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 CFEKNEA 90
Db 132 CFEKNEA 138

RESULT 11
US-08-336-728A-56
Sequence 56, Application US/08336728A
Patent No. 6207802
GENERAL INFORMATION:
APPLICANT: Zeebo, Krzyszina M.
APPLICANT: Boeselmann, Robert A.
APPLICANT: Suggs, Sidney V.
APPLICANT: Martin, Francis H.
TITLE OF INVENTION: Stem Cell Factor
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,728A
FILING DATE: 09-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/982,255
FILING DATE: 25-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/589,701
FILING DATE: 01-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/573,616
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/537,198
FILING DATE: 11-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/422,383
FILING DATE: 16-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/32956
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 282 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-728A-56
Query Match 3.6%; Score 7; DB 3; Length 282;
Best Local Similarity 100.0%; Pred. No. 53;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 178 IKENGHL 184
Db 122 IKENGHL 128

RESULT 12
US-09-328-352-6119
Sequence 6119, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6119
LENGTH: 432
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-6119
Query Match 3.6%; Score 7; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 134 GYIPINT 140
Db 268 GYIPINT 274

RESULT 13
US-08-441-139-18
Sequence 18, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltrop, Dr. Karl D.
APPLICANT: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:

LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-18

Query Match 3.6%; Score 7; DB 1; Length 504;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 177 PIRENGH 183
Db 461 PIRENGH 467

RESULT 14
US-09-453-702B-257
Sequence 257; Application US/09453702B
Patent No. 6365723

GENERAL INFORMATION:

APPLICANT: Blatner, Frederick R.
Perna, Nicole T.
Punkett, Guy
Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI

COUNTRY: US
ZIP: 53701-2113

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B

FILING DATE: 03-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955

FILING DATE: 04-DEC-1998

ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296,95017

TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 257:
SEQUENCE CHARACTERISTICS:

LENGTH: 3169 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 257:
US-09-453-702B-257

Query Match 3.6%; Score 7; DB 4; Length 3169;
Best Local Similarity 100.0%; Pred. No. 4,8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 SGILKNI 37
Db 314 SGILKNI 320

RESULT 15

US-08-817-832B-26
Sequence 26; Application US/08817832B
Patent No. 6579691

GENERAL INFORMATION:

APPLICANT: MANDELKOW, Eckhard, et al.
TITLE OF INVENTION: No. 6579691el Protein Kinase (NPK-110)

NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 S. Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois

COUNTRY: US
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,832B

FILING DATE: 28-Apr-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP95/04258

FILING DATE: 30-OCT-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94 11 7122.5

FILING DATE: 28-OCT-1994

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-817-832B-26

Query Match 3.1%; Score 6; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2,5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 LEVLNP 28
Db 1 LEVLNP 6

RESULT 16
US-08-620-151-52
Sequence 52; Application US/08620151

Patent No. 5928955

GENERAL INFORMATION:

APPLICANT: Imperial; Barbara

TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR

NUMBER OF SEQUENCES: 136

CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
CITY: Chicago
STATE: Illinois

COUNTRY: USA
ZIP: 60611-5599

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/620,151

FILING DATE: 22-MAR-1996

CLASSIFICATION: 422

ATTORNEY/AGENT INFORMATION:
NAME: Shannon, Karen L.
REGISTRATION NUMBER: 36,675
REFERENCE/DOCKET NUMBER: 8597/6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-620-151-52

Query Match 3.1%; Score 6; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKH 175
DB 15 SSLIKH 20

RESULT 17
US-08-118-270-186
Sequence 186, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-186

Query Match 3.1%; Score 6; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFFFL 8
DB 9 IFFFL 14

RESULT 18
US-08-118-270-188
Sequence 188, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-188

Query Match 3.1%; Score 6; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFFFL 8
DB 12 IFFFL 17

RESULT 19
PCT-US93-08528-186
Sequence 186, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-186

Query Match 3.1%; Score 6; DB 5; Length 32;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFFFL 8
Db 9 IFFFL 14

RESULT 20
PCT-US93-08528-188
Sequence 188, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633

INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-188

Query Match 3.1%; Score 6; DB 5; Length 32;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFFFL 8
Db 12 IFFFL 17

RESULT 21
US-08-331-625A-20
Sequence 20, Application US/08331625A
Patent No. 6057436
GENERAL INFORMATION:
APPLICANT: Miller, Timothy J.
APPLICANT: Klepfer, Sharon
APPLICANT: Reed, Albert Paul
APPLICANT: Jones, Elaine V.
TITLE OF INVENTION: Canine Coronavirus S Gene and Uses
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation - Corporate
ADDRESSEE: Patents
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,625A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/880,194
FILING DATE: 08-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/698,927
FILING DATE: 13-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/613,066
FILING DATE: 14-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Schreck, Patricia A.
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: SBC H85010-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 270-5015
TELEFAX: (215) 270-5090
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-331-625A-20

Query Match 3.1%; Score 6; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 78;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 GSWLGG 57
|||||
Db 26 GSWLGG 31

RESULT 22
US-09-494-151-20
Sequence 20, Application US/09494151
Patent No. 6372224
GENERAL INFORMATION:
APPLICANT: Miller, Timothy J.
Klepper, Sharon
Reed, Albert Paul
Jones, Elaine V.
TITLE OF INVENTION: Canine Coronavirus S Gene and Uses
therefor
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation - Corporate
Patents
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/494,151
FILING DATE: 28-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/331,625
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/698,927
FILING DATE: 13-MAY-1991
APPLICATION NUMBER: US 07/613,066
FILING DATE: 14-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Schreck, Patricia A.
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: SBC H85010-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 270-5015
TELEFAX: (215) 270-5090
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-494-151-20

Query Match 3.1%; Score 6; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 GSWLGG 57
|||||
Db 26 GSWLGG 31

RESULT 23
US-08-118-270-187
Sequence 187, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:

APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-187

Query Match 3.1%; Score 6; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFFELL 8
|||||
Db 26 IFFELL 31

RESULT 24
PCT-US93-08528-187
Sequence 187, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-187

Query Match 3.1%; Score 6; DB 5; Length 38;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFFFL 8
Db 26 IFFFL 31

RESULT 25
US-09-266-764-24
Sequence 24, Application US/09266764
Patent No. 6545139
GENERAL INFORMATION:
APPLICANT: Baylor College of Medicine
TITLE OF INVENTION: Compositions and Methods For the Treatment and
Prevention of Metastatic Disorders
FILE REFERENCE: 00A146.0122
CURRENT APPLICATION NUMBER: US/09/266,764
CURRENT FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/077,934
PRIOR FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 44
TYPE: PRT
ORGANISM: Mus sp.
US-09-266-764-24

Query Match 3.1%; Score 6; DB 4; Length 44;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 YLFL 127
Db 28 YLFL 33

RESULT 26
US-08-861-476C-5
Sequence 5, Application US/08861476C
Patent No. 6447786
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/08/861,476C
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 46
TYPE: PRT

ORGANISM: Staphylococcus aureus
US-08-861-476C-5

Query Match 3.1%; Score 6; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GILKNI 37
Db 14 GILKNI 19

RESULT 27
US-09-252-991A-22890
Sequence 22890, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22890
LENGTH: 67
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22890

Query Match 3.1%; Score 6; DB 4; Length 67;
Best Local Similarity 100.0%; Pred. No. 1,3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 FSSSLI 173
Db 23 FSSSLI 28

RESULT 28
US-09-134-001C-4520
Sequence 4520, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4520
LENGTH: 68
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4520

Query Match 3.1%; Score 6; DB 4; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 KINKIS 147
Db 45 KINKIS 50

RESULT 29
US-09-205-258-1151
Sequence 1151, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: Patencin Ver. 2.0
SEQ ID NO 1151
LENGTH: 74
TYPE: PRT
ORGANISM: Homo sapiens
US-09-205-258-1151
Query Match 3.1% Score 6; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 90 AHCXLP 95
DB 34 AHCXLP 39
RESULT 30
US-09-073-297-18
Sequence 18, Application US/09073297
Patent No. 6255114
GENERAL INFORMATION:
APPLICANT: LIGHTNER, JONATHAN EDWARD
TITLE OF INVENTION: STARCH BIOSYNTHETIC ENZYMES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,297
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/852615
FILING DATE: MAY 7, 1997
ATTORNEY/AGENT INFORMATION:
NAME: MAJARIAN, WILLIAM R.
REGISTRATION NUMBER: 41,173
REFERENCE/DOCKET NUMBER: BB-1083-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-4926
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 amino acids
TYPE: amino acid

STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-073-297-18

Query Match 3.1%; Score 6; DB 3; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 KTCPEK 87
DB 60 KTCPEK 65

RESULT 31
US-09-205-258-1038
Sequence 1038, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205,258
EARLIER FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1038
LENGTH: 81
TYPE: PRT
ORGANISM: Homo sapiens
US-09-205-258-1038

Query Match 3.1%; Score 6; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ILKNIS 38
DB 20 ILKNIS 25

RESULT 32
US-09-375-140-9
Sequence 9, Application US/09375140
Patent No. 6488540
GENERAL INFORMATION:
APPLICANT: Kavanagh, T.
TITLE OF INVENTION: A NOVEL PLASTID-TARGETING NUCLEIC ACID SEQUENCE, A
TITLE OF INVENTION: NOVEL BETA-AMYLASE SEQUENCE, A STIMULUS-RESPONSIVE
FILE REFERENCE: 9341-017
CURRENT APPLICATION NUMBER: US/09/375,140
EARLIER FILING DATE: 1999-08-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 86
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-375-140-9

Query Match 3.1%; Score 6; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SSSLIK 174
|
|
|
|
|
Db 8 SSSLIK 13

RESULT 33

US-08-665-259-2
Sequence 2, Application US/08665259
Patent No. 6028173
GENERAL INFORMATION:
APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Conners, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,259
FILING DATE: 17-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IGS-9.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 90 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-665-259-2

Query Match 3.1%; Score 6; DB 3; Length 90;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ALEAVS 72
|
|
|
|
|
Db 31 ALEAVS 36

RESULT 34

US-08-762-500-2
Sequence 2, Application US/08762500
Patent No. 6030806
GENERAL INFORMATION:
APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Conners, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME

NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,500
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,259
FILING DATE: 17-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10469
FILING DATE: 17-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IGS-9.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 90 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-762-500-2

Query Match 3.1%; Score 6; DB 3; Length 90;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ALEAVS 72
|
|
|
|
|
Db 31 ALEAVS 36

RESULT 35

US-08-936-165A-393
Sequence 393, Application US/08936165A
Patent No. 6348582
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936.165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 393:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-393

Query Match
Best Local Similarity 3.1%; Score 6; DB 4; Length 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ILKNIS 38
DB 21 ILKNIS 26

RESULT 36
US-08-331-625A-9
Sequence 9, Application US/08331625A
Patent No. 6057436
GENERAL INFORMATION:
APPLICANT: Miller, Timothy J.
APPLICANT: Klepfer, Sharon
APPLICANT: Reed, Albert Paul
APPLICANT: Jones, Elaine V.
TITLE OF INVENTION: Canine Coronavirus S Gene and Uses
TITLE OF INVENTION: Therefor
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation - Corporate
ADDRESSEE: Patents
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331.625A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/880,194
FILING DATE: 08-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/698,927
FILING DATE: 13-MAY-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/613,066
FILING DATE: 14-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Schreck, Patricia A.
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: SBC H85010-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 270-5015
TELEFAX: (215) 270-5090
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-331-625A-9

Query Match
Best Local Similarity 3.1%; Score 6; DB 3; Length 126;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 GSWLGG 57
DB 42 GSWLGG 47

RESULT 37
US-09-494-151-9
Sequence 9, Application US/09494151
Patent No. 6372224
GENERAL INFORMATION:
APPLICANT: Miller, Timothy J.
APPLICANT: Klepfer, Sharon
APPLICANT: Reed, Albert Paul
APPLICANT: Jones, Elaine V.
TITLE OF INVENTION: Canine Coronavirus S Gene and Uses
TITLE OF INVENTION: Therefor
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation - Corporate
ADDRESSEE: Patents
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/494,151
FILING DATE: 28-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/331,625
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/698,927
FILING DATE: 13-MAY-1991
APPLICATION NUMBER: US 07/613,066
FILING DATE: 14-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Schreck, Patricia A.
REGISTRATION NUMBER: 33,777
REFERENCE/DOCKET NUMBER: SBC H85010-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 270-5015
TELEFAX: (215) 270-5090
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids

TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 9
US-09-494-151-9

Query Match 3.1%; Score 6; DB 4; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 GSWLGG 57
DB 42 GSWLGG 47

RESULT 38
US-08-905-223-310
Sequence 310, Application US/08905223

GENERAL INFORMATION:
PATENT NO. 6222029
APPLICANT: Edwards, Jean-Baptiste D.
APPLICANT: Duclert, Aymeric
APPLICANT: Lacroix, Bruno
TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 503
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 310:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: AMINO ACID
TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:

NAME/KEY: s1g_peptide
LOCATION: -118...-1
IDENTIFICATION METHOD: Von Heljine matrix
OTHER INFORMATION: score 6.7
OTHER INFORMATION: seq LPAKLFGLTSA/RR
US-08-905-223-310

Query Match 3.1%; Score 6; DB 3; Length 127;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFFFL 8
DB 90 IFFFL 95

RESULT 39
US-08-858-207A-314
Sequence 314, Application US/08858207A
Patent No. 6348328

GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Nicholas, Richard
APPLICANT: Scodola, Robert
TITLE OF INVENTION: No. 6348328e1 Compounds
NUMBER OF SEQUENCES: 352
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858,207A
FILING DATE: 09-MAY-1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/017670
FILING DATE: 14-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50475
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:

INFORMATION FOR SEQ ID NO: 314:
SEQUENCE CHARACTERISTICS:
LENGTH: 128 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6348328e

US-08-858-207A-314

Query Match 3.1%; Score 6; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 LGLING 52
DB 51 LGLING 56

RESULT 40
US-09-252-991A-32648
Sequence 32648, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27

Query Match 3.1%; Score 6; DB 3; Length 127;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32648
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32648

Query Match 3.1%; Score 6; DB 4; Length 141;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NSISTI 159
|||
Db 47 NSISTI 52

RESULT 41
US-08-788-584-1
; Sequence 1, Application US/08788584
; Patent No. 5837493
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Petithory, Joanne R.
; TITLE OF INVENTION: NOVEL HUMAN GALECTINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,584
; FILING DATE: Filed Herewith
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0192 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-788-584-1

Query Match 3.1%; Score 6; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 EAHCLK 94
|||
Db 103 EAHCLK 108

RESULT 42
US-08-788-584-3

; Sequence 3, Application US/08788584
; Patent No. 5837493
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Petithory, Joanne R.
; TITLE OF INVENTION: NOVEL HUMAN GALECTINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,584
; FILING DATE: Filed Herewith
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0192 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-788-584-3

Query Match 3.1%; Score 6; DB 2; Length 149;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 EAHCLK 94
|||
Db 107 EAHCLK 112

RESULT 43
US-09-198-452A-507
; Sequence 507, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 507
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-507

Query Match 3.1%; Score 6; DB 4; Length 155;

Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 147 SLESSG 152
DB 94 SLESSG 99

RESULT 44

US-09-162-021B-10
Sequence 10, Application US/09162021B
Patent No. 6337391
GENERAL INFORMATION:
APPLICANT: H. William Harris
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
TITLE OF INVENTION: Polyclonal-Sensing Receptor in Aquatic
FILE REFERENCE: 2856.1001-007
CURRENT APPLICATION NUMBER: US/09/162.021B
PRIOR FILING DATE: 1998-09-28
PRIOR APPLICATION NUMBER: PCT/US97/05031
PRIOR FILING DATE: 1997-03-27
PRIOR APPLICATION NUMBER: 08/622,738
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 157
TYPE: PRT
ORGANISM: parailchthys dentatus
US-09-162-021B-10

Query Match 3.1%; Score 6; DB 4; Length 157;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 FSSSLI 173
DB 56 FSSSLI 61

RESULT 45

US-08-484-993B-45
Sequence 45, Application US/08484993B
Patent No. 5837497
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,993B
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341

FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-993B-45

Query Match 3.1%; Score 6; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 SSGENS 155
DB 101 SSGENS 106

RESULT 46

US-08-484-158B-45
Sequence 45, Application US/08484158B
Patent No. 5976545
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Pharmaceutical Compositions for
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,158B
FILING DATE: 07-JUNE-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 09-NOV-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-92
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:

LENGTH: 164 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-158B-45

Query Match 3.1%; Score 6; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 SGENS 155
Db 101 SGENS 106

RESULT 47
US-08-484-596A-45
Sequence 45, Application US/08484596A
Patent No. 5981228
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunoprotection
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,596A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 11-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-596A-45

Query Match 3.1%; Score 6; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 SGENS 155
Db 101 SGENS 106

RESULT 48

US-08-480-150A-45
Sequence 45, Application US/08480150A
Patent No. 5989550
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunoprotection
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,150A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/149,223
FILING DATE: 09-NOV-1993
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-150A-45

Query Match 3.1%; Score 6; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 SGENS 155
Db 101 SGENS 106

RESULT 49
US-08-458-731-45
Sequence 45, Application US/08458731
Patent No. 6001599
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunoprotection
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois

COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,731
FILING DATE: 09-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-731-45

Query Match 3.1%; Score 6; DB 3; Length 164;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 SSGENS 155
DB 101 SSGENS 106

RESULT 50
US-08-149-223A-45
Sequence 45, Application US/08149223A
Patent No. 6027727
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunoccontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,223A
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341

FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-149-223A-45

Query Match 3.1%; Score 6; DB 3; Length 164;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 SSGENS 155
DB 101 SSGENS 106

Search completed: January 28, 2004, 18:24:24
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using BW model

Run on: January 28, 2004, 16:03:50 ; Search time 69 Seconds
(without alignments)
443,974 Million cell updates/sec

Title: US-09-881-556A-4

Perfect score: 193
Sequence: 1 HEIFFFLPLPISIFPDSNDA.....KHAPIKGNHNNOSKYSK 193

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size: 5

Total number of hits satisfying chosen parameters: 18238

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database:

1: A.Geneseq_19Jun03.*
2: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1986.DAT.*
9: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1987.DAT.*
10: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1989.DAT.*
12: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1990.DAT.*
13: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1991.DAT.*
14: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1992.DAT.*
15: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1993.DAT.*
16: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1994.DAT.*
17: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1995.DAT.*
18: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1996.DAT.*
19: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1997.DAT.*
20: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1998.DAT.*
21: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1999.DAT.*
22: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2000.DAT.*
23: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2001.DAT.*
24: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2002.DAT.*
25: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	193	100.0	193	23	AA50658
2	9	4.7	19938	24	ABP76680
3	7	3.6	72	20	AA36135
4	7	3.6	72	20	AA36182
5	7	3.6	75	23	ABP08491
6	7	3.6	75	24	ABUS5703
7	7	3.6	81	22	ABG07587
8	7	3.6	81	22	ABG18998
9	7	3.6	84	21	AA601515

10	7	3.6	100	19	AAW75205	Human secreted pro
11	7	3.6	100	21	AAE43362	Human ORFX ORF3126
12	7	3.6	100	23	AAE26968	Human gene 10 enco
13	7	3.6	100	23	AAE27106	Human gene 10 enco
14	7	3.6	100	24	ABU64979	Human secreted pro
15	7	3.6	101	21	AAV65396	Human 5' EST relat
16	7	3.6	106	20	AAV36137	Human secreted pro
17	7	3.6	106	20	AAV36184	Human secreted pro
18	7	3.6	106	20	AAV35987	Extended human sec
19	7	3.6	106	23	ABP97399	Novel human protei
20	7	3.6	115	23	ABG60021	Human DTRP polype
21	7	3.6	128	24	ABR41138	Human DTRP enzyme
22	7	3.6	139	22	ABG08306	Novel human diageno
23	7	3.6	140	23	ABE47391	Lieteria monocytog
24	7	3.6	146	23	ABG63138	Bacillus glyphosat
25	7	3.6	146	23	ABG63139	Bacillus glyphosat
26	7	3.6	146	23	ABG63143	Bacillus glyphosat
27	7	3.6	146	23	ABG63150	Bacillus glyphosat
28	7	3.6	146	23	ABG63156	Bacillus glyphosat
29	7	3.6	146	23	ABG63160	Bacillus glyphosat
30	7	3.6	146	23	ABG63162	Bacillus glyphosat
31	7	3.6	146	23	ABG63163	Bacillus glyphosat
32	7	3.6	146	23	ABG63168	Bacillus glyphosat
33	7	3.6	146	23	ABG63171	Bacillus glyphosat
34	7	3.6	146	23	ABG63183	Bacillus glyphosat
35	7	3.6	146	23	ABG63185	Bacillus glyphosat
36	7	3.6	146	23	ABG63186	Bacillus glyphosat
37	7	3.6	146	23	ABG63187	Bacillus glyphosat
38	7	3.6	146	23	ABG63188	Bacillus glyphosat
39	7	3.6	146	23	ABG63189	Bacillus glyphosat
40	7	3.6	146	23	ABG63190	Bacillus glyphosat
41	7	3.6	146	23	ABG63193	Bacillus glyphosat
42	7	3.6	146	23	ABG63201	Bacillus glyphosat
43	7	3.6	146	23	ABG63202	Bacillus glyphosat
44	7	3.6	146	23	ABG63204	Bacillus glyphosat
45	7	3.6	146	23	AAU16056	Human novel secret
46	7	3.6	151	24	AAU55125	Human novel polype
47	7	3.6	152	21	AAU16890	Zee may's protein f
48	7	3.6	160	23	ABG72337	Human hypoxin indu
49	7	3.6	162	22	ABB11250	Human Fc-gamma rec
50	7	3.6	163	21	AAU19020	Zee may's protein f
51	7	3.6	172	20	AAV27240	Amino acid sequenc
52	7	3.6	176	23	ABP30553	Streptococcus poly
53	7	3.6	199	21	AAU64633	Arabidopsis thalia
54	7	3.6	204	17	AAU14318	(His)6 modified ch
55	7	3.6	206	23	ABG60039	Human DTRP polype
56	7	3.6	212	24	AAU85558	Clone #59303 (L142
57	7	3.6	212	24	AAU66930	Human lung cancer-
58	7	3.6	212	24	AAU66433	Lung cancer therap
59	7	3.6	212	24	AAU56468	Lung cancer-associ
60	7	3.6	212	24	AAU07497	Protein differenti
61	7	3.6	212	24	AAU03524	Angiogenesis-asso
62	7	3.6	212	24	AAU78599	Human protein SEQ
63	7	3.6	223	22	ABP27043	Streptococcus poly
64	7	3.6	223	22	AAW79583	Human protein SEQ
65	7	3.6	242	20	AAV27241	Amino acid sequenc
66	7	3.6	242	20	AAV27242	Amino acid sequenc
67	7	3.6	246	21	AAU41678	Arabidopsis thalia
68	7	3.6	257	23	ABP26056	Streptococcus poly
69	7	3.6	270	21	AAU41677	Arabidopsis thalia
70	7	3.6	282	22	AAU96949	Chicken stem cell
71	7	3.6	294	22	AAU92566	C glutamicum prote
72	7	3.6	294	22	AAU86505	C. glutamicum sucd
73	7	3.6	294	22	AAU79525	Corynebacterium gl
74	7	3.6	308	22	AAU59299	Drosophila melanog
75	7	3.6	315	22	AAU78762	Drosophila melanog
76	7	3.6	342	22	AAU78774	Human protein sequ
77	7	3.6	344	22	ABG18999	Human protein sequ
78	7	3.6	357	21	AAU26404	Novel human diageno
79	7	3.6	397	21	AAU26443	Drosophila melanog
80	7	3.6	397	21	AAU20916	Drosophila melanog
81	7	3.6	397	21	AAU20916	Drosophila odorant
82	7	3.6	397	22	ABW71190	Drosophila melanog

83	7	3.6	429	22	AA091564	C glutamicum prote
84	7	3.6	488	22	AB027955	Novel human diagno
85	7	3.6	489	22	AB050238	Human transcriptio
86	7	3.6	489	23	AA093047	Arabidopsis thaliana
87	7	3.6	549	24	AA016426	Human nucleic acid
88	7	3.6	577	22	AB060273	Drosophila melanog
89	7	3.6	627	23	AB077914	High level promote
90	7	3.6	754	22	AB065145	Drosophila melanog
91	7	3.6	851	22	AB07753	Novel human diagno
92	7	3.6	903	22	AB065230	Drosophila melanog
93	7	3.6	1297	23	AA096927	Haemophilus influe
94	7	3.6	1297	23	AA096930	Haemophilus influe
95	7	3.6	1298	23	AA096928	Haemophilus influe
96	7	3.6	1298	23	AA096929	Haemophilus influe
97	6	3.1	8	17	AA098246	Pig brain MARK lys
98	6	3.1	14	22	AA011205	Parvovirus epitope
99	6	3.1	15	22	AA011204	Parvovirus epitope
100	6	3.1	18	22	AA011203	Parvovirus epitope

ALIGNMENTS

RESULT 1
AAMS0658 standard; Protein; 193 AA.

AC AAMS0658;
DT 04-APR-2002 (first entry)
DE Cryosporidium parvum antigenic polypeptide 2 (Ag2).
XX
XX
XX AG1, antigen; immunogen; cryosporidiosis; infection; diagnosis;
XX therapy; vaccine.
XX
XX Cryosporidium parvum.
XX
XX
FH Key Location/Qualifiers
FT Modified-site 36..38 /note= "asn is N-glycosylated"
FT Modified-site 51..53 /note= "asn is N-glycosylated"
FT
XX
XX MO200196370-A2.
XX
XX 20-DEC-2001.
XX
XX 15-JUN-2001; 2001MO-CA00856.
XX
XX 15-JUN-2000; 2000US-212083P.
XX
XX (UYMA-) UNIV MANITOBA.
XX
XX Simonsen JN;
XX
XX WPI: 2002-139704/18.
XX N-PSDB; ABA91278.
XX
XX
XX Novel isolated nucleic acid comprising coding sequence for an
XX immunogenic Cryosporidium parvum antigenic polypeptide useful as
XX vaccine components for treating or preventing C. parvum infection in
XX mammals
XX
XX Claim 9; Fig 2A; 61pp; English.
XX
XX
XX The present sequence comprises immunogenic Cryosporidium parvum
XX antigenic polypeptide 2 (Ag2), as predicted from the incomplete
XX open reading frame given in ABA91278. The predicted Ap2 protein
XX has a mol.wt. of 21.8 KDa and a pI of 6.23. The invention relates
XX to the discovery of novel genes encoding C. parvum antigenic
XX polypeptides Ag1 and Ag2, the characterisation of these proteins,
XX and antibodies that recognise epitopes of Ap1 and Ap2. Also

CC provided are recombinant vectors, host cells and methods of
CC producing recombinant C. parvum antigenic proteins. The proteins,
CC their immunogenic fragments and analogues, and/or chimeric proteins,
CC can be used either alone or in combination with other C. parvum
CC antigens in novel subunit vaccines to provide protection from
CC cryosporidial infection in mammals. Antibodies generated against
CC the proteins can be used as therapeutic agents. The antigens and
CC antibodies can also be used as diagnostics of infection.
XX
SQ Sequence 193 AA;

Query Match 100.0%; Score 193; DB 23; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.7e-187;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HEIFFLPISISFPDSNDKSLFVNPDSSGILKIKSTFEIKFELGLNSWLGDI 60
DB 1 HEIFFLPISISFPDSNDKSLFVNPDSSGILKIKSTFEIKFELGLNSWLGDI 60
QY 61 ILDRKALAVSYSIACVFYTKTCFEKNEAHCLEPRAENKMTFGEKDLATTLQSSNS 120
DB 61 ILDRKALAVSYSIACVFYTKTCFEKNEAHCLEPRAENKMTFGEKDLATTLQSSNS 120
QY 121 EYLLFTWNNCIGYIPINTNKINKISLESGENSISTIGWSIIDGFSSSLIKAPIKE 180
DB 121 EYLLFTWNNCIGYIPINTNKINKISLESGENSISTIGWSIIDGFSSSLIKAPIKE 180
QY 181 NGHLNNOESKYSK 193
DB 181 NGHLNNOESKYSK 193

RESULT 2

ID ABP76680 standard; Protein; 19938 AA.

AC ABP76680;
DT 26-FEB-2003 (first entry)
DE Streptomyces viridochromogenes Avi gene cluster polypeptide frame 4.
XX
XX
XX Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;
XX medicine; Staphylococcus aureus; biosynthetic gene cluster; enzyme.
XX
XX Streptomyces viridochromogenes.
XX
XX
XX WO200268436-A1.
XX
XX 06-SEP-2002.
XX
XX 24-AUG-2001; 2001MO-EP09815.
XX
XX 25-FEB-2001; 2001DE-1009166.
XX
XX (COMB-) COMBINATURE BIOPHARM AG.
XX
XX Weitnauer G, Muehlenweg A, Trefzer A, Bechtold A;
XX
XX WPI: 2003-018650/01.
XX N-PSDB; AB237516.
XX
XX
XX New avilamycin derivatives, useful for treatment of infections, and
XX nucleic acid encoding avilamycin synthesis enzymes
XX
XX
XX Example 1; Page 68-301; 31pp; German.
XX
XX
XX The invention relates to avilamycin derivatives (I) with antibacterial,
XX virucide, protozoacide and fungicide activity. (I) are useful for
XX treatment of infections (bacterial, viral, protozoal or fungal), in human
XX or veterinary medicine, particularly where caused by Staphylococcus
XX aureus. (I) are more hydrophilic than known avilamycins. The present
XX sequence is that of an avilamycin synthesis enzyme from the Streptomyces

CC viridochromogenes Avilamycin A biosynthetic gene cluster
 CC (AB237515-AB237516).
 XX
 SO Sequence 19938 AA;

Query Match 4.7%; Score 9; DB 24; Length 19938;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 150 SSGENIST 158
 DB 19809 SSGENIST 19817

RESULT 3
 AAY36135
 ID AAY36135 standard; protein; 72 AA.

XX AAY36135;
 AC AAY36135;
 DT 23-SEP-1999 (first entry)
 XX Human secreted protein #7.
 DE Human secreted protein #7.
 XX
 KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;
 KM diagnostic; gene therapy; chromosome mapping; secretion vector.
 XX
 OS Homo sapiens.

XX WO9925825-A2.
 PN 27-MAY-1999.
 PD 13-NOV-1998; 98WO-IB01862.
 XX

PR 04-SEP-1998; 98US-0099273.
 PR 13-NOV-1997; 97US-0066677.
 PR 17-DEC-1997; 97US-0069957.
 PR 09-FEB-1998; 98US-0074121.
 PR 13-APR-1998; 98US-0081563.
 PR 10-AUG-1998; 98US-0096116.

XX (GEST) GENSET.
 PA Bougueleret L, Duclet A, Dumas Milne Edwards J;
 PI WPI; 1999-347472/29.
 DR N-PSDB; AAX97819.
 XX

XX Extended cDNAs encoding secreted proteins
 PT Example 28; Page 235; 307pp; English.

CC AAY36129-Y36222 represent novel human secreted proteins encoded by the
 CC extended cDNA sequences represented in AAX97813-X97906. The proteins
 CC of the invention have cytosolic, thrombotic and osteopathic activity.
 CC The extended cDNAs can be used to express secreted proteins or parts of
 CC them or to obtain antibodies capable of binding to the secreted
 CC proteins. They may also be used in diagnostic, forensic, gene therapy
 CC and chromosome mapping procedures. Uses also include design of
 CC expression vectors and secretion vectors.

XX Sequence 72 AA;

Query Match 3.6%; Score 7; DB 20; Length 72;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 NKISLES 150
 DB 59 NKISLES 65

RESULT 4
 AAY36182
 ID AAY36182 standard; protein; 72 AA.
 XX

AC AAY36182;
 XX 23-SEP-1999 (first entry)
 DT Human secreted protein #54.
 XX

KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;
 KM diagnostic; gene therapy; chromosome mapping; secretion vector.

XX Homo sapiens.
 OS WO9925825-A2.
 PN 27-MAY-1999.
 PD 13-NOV-1998; 98WO-IB01862.
 XX

PR 04-SEP-1998; 98US-0099273.
 PR 13-NOV-1997; 97US-0066677.
 PR 17-DEC-1997; 97US-0069957.
 PR 09-FEB-1998; 98US-0074121.
 PR 13-APR-1998; 98US-0081563.
 PR 10-AUG-1998; 98US-0096116.

XX (GEST) GENSET.
 PA Bougueleret L, Duclet A, Dumas Milne Edwards J;
 PI WPI; 1999-347472/29.
 DR N-PSDB; AAX97866.
 XX

XX Extended cDNAs encoding secreted proteins
 PT Claim 7; Page 289-290; 307pp; English.

CC AAY36129-Y36222 represent novel human secreted proteins encoded by the
 CC extended cDNA sequences represented in AAX97813-X97906. The proteins
 CC of the invention have cytosolic, thrombotic and osteopathic activity.
 CC The extended cDNAs can be used to express secreted proteins or parts of
 CC them or to obtain antibodies capable of binding to the secreted
 CC proteins. They may also be used in diagnostic, forensic, gene therapy
 CC and chromosome mapping procedures. Uses also include design of
 CC expression vectors and secretion vectors.

XX Sequence 72 AA;

Query Match 3.6%; Score 7; DB 20; Length 72;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 NKISLES 150
 DB 59 NKISLES 65

RESULT 5
 ABP08491
 ID ABP08491 standard; Protein; 75 AA.
 XX

AC ABP08491;
 XX 24-JUN-2002 (first entry)
 DT Human ORFX protein sequence SEQ ID NO:16964.
 XX

KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;

KM hypertension; hypothyroidism; cholesterol ester storage disease;
 KM immune deficiency; immune disorder; infectious disease;
 KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KM myasthenia gravis.
 OS Homo sapiens.
 XX MO200192523-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 29-MAY-2001; 2001WO-US10836.
 XX
 XX 30-MAY-2000; 2000US-206132P.
 XX 29-AUG-2000; 2000US-228716P.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Shimkets RA, Leach MD;
 XX
 XX WPI: 2002-106308/14.
 XX N-PSDB; ABN24243.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 XX
 PS Disclosure; SEQ ID 16964; 1037pp; English.
 XX
 XX The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP1500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumors, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pot_sequences.
 XX
 SQ Sequence 75 AA;
 QY
 DB Best Local Similarity 3.6%; Score 7; DB 23; Length 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 ILKNIST 39
 DB 49 ILKNIST 55
 RESULT 6
 ABUS5703
 ID ABUS5703 standard; Protein; 75 AA.
 XX
 AC ABUS5703;
 XX
 DT 17-MAR-2003 (first entry)

XX
 DE Rice growth regulating protein, OSGREP4.
 XX
 XX ss; plant; growth regulating protein; PSK; phytoosulphokine-alpha;
 KM transgenic; marker-assisted breeding; agriculture; horticulture.
 XX
 XX Oryza sativa.
 XX
 XX MO200283901-A2.
 XX
 XX 24-OCT-2002.
 XX
 XX 11-APR-2002; 2002WO-EP04035.
 XX
 XX 12-APR-2001; 2001US-283313P.
 XX
 XX (CROP-) CROPDESIGN NV.
 XX (SAUT/) SAUTER M.
 XX
 XX Sauter M, Lorbiecke R, Mironov V, Frankard V, Dillen W, Lejeune P;
 XX Broekaert W;
 XX
 XX WPI: 2003-093023/08.
 XX N-PSDB; ABX74171, ABX74172.
 XX
 XX New isolated GREP nucleic acids and polypeptides controlling or
 PT modifying the growth characteristics of a plant, useful for a wide
 PT application in agricultural and horticultural practices, and in vitro
 PT plant cell and tissue culture -
 XX
 XX Claim 5; Fig 22; 156pp; English.
 XX
 XX The invention relates to a new isolated nucleic acid encoding a growth
 CC regulating protein (GREP), where the protein conforms to the GREP
 CC signature sequence appearing as ABUS5722, and the GREP proteins.
 CC Also included are a vector comprising a GREP nucleic acid, or a vector
 CC encoding the rice growth regulating polypeptide OSGREP (phytoosulphokine-
 CC alpha) where the growth regulating proteins regulate growth and/or
 CC development response in intact plants, a GREP transgenic plant,
 CC a transgenic plant comprising the vector, seed/pollen and a harvestable
 CC part of propagation material from the transgenic plant, a host cell
 CC comprising the GREP nucleic acid, an antisense molecule consisting of
 CC 14-100 nucleotides targeted to the GREP signature nucleotide
 CC sequence appearing as ABX74197, an anti-GREP antibody which specifically
 CC recognises a GREP or its fragment, a method for altering growth and/or
 CC activity of a plant or plant cell/storage organ comprising modulating the
 CC level and/or activity of a GREP or the rice growth regulating peptide
 CC OSGP, a method for downregulating levels of a GREP or OSGP gene product
 CC or its activity (comprising administering or exposing GREP or OSGP
 CC antibodies or gene products to cells, tissues or organs of a plant)
 CC a method for identifying an allele with desired features of a gene
 CC encoding a GREP which comprises isolating alleles for a GREP and
 CC testing the features of the allele by expression in a transgenic plant,
 CC a method for identifying an allele of GREP and selecting an allele
 CC with desired features which comprises the use of genes encoding GREP,
 CC or sequences located in the genome in the neighbourhood of GREP genes,
 CC as molecular markers for different GREP alleles and selecting specific
 CC GREP alleles by marker-assisted breeding and a method for identifying
 CC regulatory sequences of GREP growth regulating polypeptide genes.
 CC The methods and compositions are for controlling or modifying the
 CC growth characteristics of a plant or its organs and tissues, useful for
 CC in vitro plant cell and tissue culture. The present sequence is a
 CC growth regulating protein.
 XX
 SQ Sequence 75 AA;
 QY
 DB Best Local Similarity 3.6%; Score 7; DB 24; Length 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 LRPISIS 13
 DT

DB 14 LEPISIS 20

RESULT 7

ABG07587

ID ABG07587 standard; Protein; 81 AA.

XX

AC ABG07587;

XX

DT 13-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #7578.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN MO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001MO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

XX

PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

XX

DR N-PSDB; AAS71774.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensic, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX

PS Claim 20; SEQ ID NO 37946; 103pp; English.

XX

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensic, gene mapping, identification of mutations in

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

XX Sequence 81 AA;

SO

Query Match 3.6%; Score 7; DB 22; Length 81;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 CFEKNEA 90

DB 25 CFEKNEA 31

RESULT 8

ABG18998

ID ABG18998 standard; Protein; 81 AA.

XX

AC ABG18998;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #18989.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN MO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001MO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

XX

PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

XX

DR N-PSDB; AAS83185.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensic, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX

PS Claim 20; SEQ ID NO 49357; 103pp; English.

XX

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensic, gene mapping, identification of mutations in

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

XX Sequence 81 AA;

SO

Query Match 3.6%; Score 7; DB 22; Length 81;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 CFEKNEA 90

DB 25 CFEKNEA 31

RESULT 9

AA01515
ID AAG01515 standard; Protein: 84 AA.
XX
AC AAG01515;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 5596.
XX
KW Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI: 2000-500381/45.
XX
DR N-PSDB; AAC01521.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX
PS Claim 13; SEQ ID 5596; 71bp + CD-ROM; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
CC
XX
SQ Sequence 84 AA;
XX
XX
Query Match 3.6%; Score 7; DB 21; Length 84;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 144 NKISLES 150
Db 35 NKISLES 41

XX
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KM inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9840483-A2.
XX
PD 17-SEP-1998.
XX
PF 12-MAR-1998; 98WO-US04858.
XX
XX 19-DEC-1997; 97US-0068368.
XX
PR 14-MAR-1997; 97US-0040710.
XX
PR 14-MAR-1997; 97US-0040762.
XX
PR 30-MAY-1997; 97US-0048100.
XX
PR 30-MAY-1997; 97US-0048189.
XX
PR 30-MAY-1997; 97US-0048357.
XX
PR 30-MAY-1997; 97US-0050934.
XX
PR 06-JUN-1997; 97US-0048970.
XX
PR 05-SEP-1997; 97US-0057765.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ferrie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;
XX Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;
PI Wei YF, Young PE, Zeng Z;
XX
XX WPI: 1998-520811/44.
XX
DR N-PSDB; AAV34295.
XX
XX Isolated human poly(nucleotide(s) encoding secretory peptide(s) -
PT used to develop products for the diagnosis and treatment of e.g.
PT inflammation, cancers, CNS disorders or immune system disorders
XX
XX
PS Claim 1; Page 159; 201bp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AAV3427) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 28 novel genes and their fragments (nucleic
CC acid sequences: AAV34286-V34325; amino acid sequences AAW75196-W75235)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 28
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAV34286 for described uses).
XX
XX
SQ Sequence 100 AA;
XX
XX
Query Match 3.6%; Score 7; DB 19; Length 100;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 144 NKISLES 150
Db 53 NKISLES 59

RESULT 10
AAW75205
ID AAW75205 standard; Protein: 100 AA.
XX
AC AAW75205;
XX
DT 29-JAN-1999 (first entry)
XX
DE Human secreted protein encoded by gene 10 clone HUTA17.
XX
XX Human, secreted protein; fusion protein; gene therapy; protein therapy;
KM diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KM developmental abnormality; foetal deficiency; blood; allergy; renal;
XX

RESULT 11
AAB43362
ID AAB43362 standard; Protein: 100 AA.
XX
AC AAB43362;
XX
DT 08-FEB-2001 (first entry)
XX

DE Human ORFX ORF3126 polypeptide sequence SEQ ID NO:6252.
 XX
 KM Human: open reading frame: ORFX: detection; cytostatic; hepatotropic;
 KM vulnery; antipariatic; antiparkinsonian; nootropic; neuroprotective;
 KM anticonvulsant; osteopathic; anticholinergic; immunosuppressant; cardiac;
 KM immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antihypertensive;
 KM antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;
 KM antineoplastic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesteryl ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KM chromoblasts; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN MO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000MO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkete RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC7571.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 5440; 5507pp; English.
 XX
 CC AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipariatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; anticholinergic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotrophic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineoplastic;
 CC antihypertensive; and antineoplastic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesteryl ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergic, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 100 AA;

Query Match 3.6%; Score 7; DB 21; Length 100;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

151 SGENSIS 157
 |||||

DB 28 SGENSIS 34
 RESULT 12
 AA26968
 ID AA26968 standard; Protein, 100 AA.
 XX
 AC AA26968;
 XX
 DT 13-DEC-2002 (first entry)
 XX
 DE Human gene 10 encoded secreted protein HTPA17, SEQ ID NO:61.
 XX
 KM Human: immunodeficiency; X-linked agammaglobulinaemia; septic shock;
 KM autoimmune disorder; rheumatoid arthritis; multiple sclerosis; cancer;
 KM Grave's disease; diabetes mellitus; haematopoietic disorder; stroke;
 KM respiratory disorder; asthma; allergy; gastrointestinal disorder;
 KM inflammatory bowel disease; neurodegenerative disorder; hepatitis;
 KM Parkinson's disease; Alzheimer's disease; cardiovascular disorder;
 KM atherosclerosis; myocarditis; renal disorder; fungicide; virucide;
 KM hyperproliferative disorder; acute glomerulonephritis; tonsillitis;
 KM respiratory disorder; rhinitis; sinusitis; neurological disease;
 KM endocrine disorder; Addison's disease; reproductive system disorder;
 KM endometriosis; vasotrophic; vulnery; cytostatic; nootropic; cardiac;
 KM anti-HIV; tranquilliser; gout; antiparasitic; chromosome 14.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..32
 FT Protein /label= Signal_peptide
 FT 33..100
 FT /note= "Human mature secreted protein"
 XX
 PN US2002077287-A1.
 XX
 PD 20-JUN-2002.
 XX
 PF 11-MAY-2001; 2001US-0852659.
 XX
 PR 11-SEP-1998; 98US-0152060.
 XX
 PA (RUBE/) RUBEN S M.
 PA (ROSE/) ROSEN C A.
 PA (LIYY/) LI Y.
 PA (ZENG/) ZENG Z.
 PA (KIYAN/) KIYAN H.
 PA (FISC/) FISCHER C L.
 PA (LIHH/) LI H.
 PA (SOPP/) SOPPET D R.
 PA (GENT/) GENTZ R L.
 PA (WEIY/) WEI Y.
 XX
 PI Ruben SM, Rosen CA, Li Y, Zeng Z, Kiyaw H, Fischer CL, Li H;
 PI Soppet DR, Gentz RL, Wei Y, Moore PA, Young PE, Greene JM;
 PI Ferrite AM;
 XX
 DR WPI: 2002-598780/64.
 DR N-PSDB; AAD44645.
 XX
 PT Novel human secreted polypeptides and polynucleotides for diagnosing,
 PT preventing, treating immune, hyperproliferative, cardiovascular,
 PT neurological, reproductive disorders and identifying modulators of
 PT therapeutic use -
 XX
 SQ Claim 11; Page 177-178; 209pp; English.
 XX
 CC AAD44636-AAD44676 represent cDNAs corresponding to 28 human secreted
 CC protein genes, and AA26969-AA26999 represent the proteins they encode.
 CC AA27000-AA27025 represent human secreted protein fragments or cDNA
 CC variants. The secreted proteins and genes are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Specific uses are described for each of the 28 genes, based

CC on the tissues in which they are most highly expressed and include
 CC developing products for the diagnosis or treatment of immunodeficiencies,
 CC e.g., X-linked agammaglobulinemia, B cell immunodeficiencies, severe
 CC combined immunodeficiencies, autoimmune disorders e.g., systemic lupus
 CC erythematosus, rheumatoid arthritis, multiple sclerosis, autoimmune
 CC thyroiditis, autoimmune haemolytic anaemia, Goodpasture's syndrome,
 CC Grave's disease, diabetes mellitus, dermatitis, inflammatory conditions
 CC including septic shock, sepsis, reperfusion injury, inflammatory bowel
 CC disease, Crohn's disease, haematopoietic disorders, respiratory
 CC disorders e.g., asthma and allergy, gastrointestinal disorders e.g.,
 CC inflammatory bowel disease, cancer e.g., gastric, ovarian, lung,
 CC liver, bladder and breast), central nervous system (CNS) disorders e.g.,
 CC ischaemic brain injury and/or stroke, neurodegenerative disorders e.g.,
 CC Parkinson's disease and Alzheimer's disease, AIDS-related dementia and
 CC prion disease, cardiovascular disorders e.g., myocarditis, arrhythmias,
 CC atherosclerosis, inflammatory disorders e.g., hepatitis, gout, trauma,
 CC pancreatitis, sarcoidosis and allogeneic transplant rejection, blood-
 CC related disorder (thrombosis, arterial thrombosis, atherosclerosis),
 CC hyperproliferative disorders, respiratory disorders e.g., rhinitis,
 CC sinusitis, tonsillitis, lung cancer, allergic disorders, pneumonitis,
 CC renal disorders, endocrine disorders e.g., hyperthyroidism, Addison's
 CC disease, hyperparathyroidism, infectious diseases and reproductive system
 CC disorders e.g., endometriosis. The present sequence represents a human
 CC secreted protein of the invention.

XX Sequence 100 AA;

Query Match 3.6%; Score 7; DB 23; Length 100;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 144 NKISLES 150
 Db 53 NKISLES 59

RESULT 13
 ID AAE27106 standard; Protein; 100 AA.

XX AAE27106;
 DT 13-DEC-2002 (first entry)

XX Human gene 10 encoded secreted protein HUTA17, SEQ ID NO:61.

XX Human; secreted protein; autoimmune disease; hyperproliferative disorder;
 KW rheumatoid arthritis; neoplasm; cerebrovascular disorder; angiogenesis;
 KW cerebral ischaemia; cardiovascular disorder; nervous system disorder;
 KW cardiac arrest; Alzheimer's disease; ocular disorder; wound healing;
 KW infection; corneal infection; skin aging; food additive; preservative;
 KW tissue regeneration; immunosuppressive; antiproliferative; cytostatic;
 KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
 KW antibacterial; virucide; fungicide; ophthalmological; gene therapy;
 KW vulnary.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Peptide 1..32
 FT /label= Signal_peptide
 FT Protein 33..100
 FT /note= "Mature human secreted protein"

XX US2002076756-A1.

XX 20-JUN-2002.

XX 11-MAY-2001; 2001US-0853161.

XX 02-FEB-2001; 2001US-265583P.

XX

PA (RUBE/) RUBEN S M.
 PA (ROSE/) ROSEN C A.
 PA (LIYY/) LI Y.
 PA (ZENG/) ZENG Z.
 PA (KYAM/) KYAM H.
 PA (FISC/) FISCHER C L.
 PA (LIHH/) LI H.
 PA (SOPP/) SOPPET D R.
 PA (GENT/) GENTZ R L.
 PA (WEIV/) WEI Y.
 PA (MOOR/) MOORE P A.
 PA (YOUN/) YOUNG P E.
 PA (GREE/) GREENE J M.
 PA (FERR/) FERRIE A M.
 XX
 PI Ruben SM, Rosen CA, Li Y, Zeng Z, Kyaw H, Fischer CL, Li H;
 PI Soppet DR, Gentz RL, Wei Y, Moore PA, Young PE, Greene JM,
 PI Ferrie AM;
 XX
 DR WPI; 2002-574454/61.
 DR N-PSDB; AAD44863.
 XX
 PT New nucleic acid molecules encoding 28 human secreted proteins, useful
 PT for diagnosing, preventing, treating or ameliorating medical conditions
 PT and as food additives or preservatives
 XX
 PS Claim 11; Page 177-178; 209pp; English.

CC AAD44854-AA044984 represent cDNAs corresponding to 28 human secreted
 CC protein genes, and AAE27097-AAE27137 represent the proteins they encode.
 CC AAE27138-AAE27164 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Secreted protein sequences of the invention are useful for the
 CC diagnosis or treatment of disorders such as autoimmune diseases (e.g.,
 CC rheumatoid arthritis), hyperproliferative disorders (e.g., neoplasms of
 CC the breast or liver), cerebrovascular disorders (e.g., cerebral ischaemia,
 CC angiogenesis), cardiovascular disorders (e.g., cardiac arrest), nervous
 CC system disorders (e.g., Alzheimer's disease), infections caused by fungi,
 CC bacteria and viruses and ocular disorders (e.g., corneal infection). The
 CC polypeptides can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, for supporting cell culture of primary tissues,
 CC to regenerate tissues and in chemotaxis. They can also be used as food
 CC additives or preservative to increase or decrease storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC and other nutritional components. The present sequence represents a human
 CC secreted protein of the invention.

XX Sequence 100 AA;

Query Match 3.6%; Score 7; DB 23; Length 100;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 144 NKISLES 150
 Db 53 NKISLES 59

RESULT 14
 ID AB064979 standard; Protein; 100 AA.

XX AB064979;

XX 15-MAY-2003 (first entry)

XX Human secreted protein gene 10, protein.

XX Secreted protein; immunodeficiency; multiple sclerosis;
 KW severe combined immunodeficiency; autoimmune disorder; cancer;
 KW rheumatoid arthritis; diabetes mellitus; haematopoietic disorder;

inflammatory condition; septic shock; inflammatory bowel disease;
 Crohn's disease; respiratory disorder; asthma; allergy; stroke;
 gastrointestinal disorder; central nervous system disorder;
 ischaemic brain injury; neurodegenerative disorder; Parkinson's disease;
 Alzheimer's disease; cardiovascular disorder; atherosclerosis;
 blood-related disorder; thrombosis; atherosclerosis; renal disorder;
 hyperproliferative disorder; acute glomerulonephritis; Addison's disease;
 endocrine disorder; liver disease; reproductive system disorder;
 endometriosis; infectious disease; pancreatic disorder; vaccine;
 wound repair; angiogenesis; lymphatic disorder; hair loss; body weight;
 body height; hair colour; human.
 Homo sapiens.
 US2002172994-A1.
 21-NOV-2002.
 11-MAY-2001; 2001US-0852797.
 14-MAR-1997; 97US-040710P.
 14-MAR-1997; 97US-040762P.
 30-MAY-1997; 97US-048100P.
 30-MAY-1997; 97US-048189P.
 30-MAY-1997; 97US-048357P.
 30-MAY-1997; 97US-050934P.
 06-JUN-1997; 97US-048970P.
 05-SEP-1997; 97US-057765P.
 19-DEC-1997; 97US-068368P.
 02-FEB-2001; 2001US-265583P.
 12-MAR-1998; 98WO-US04858.
 11-SEP-1998; 98US-0152060.
 (RUBE/) RUBEN S M.
 (ROSE/) ROSEN C A.
 (LIYY/) LI Y.
 (ZENG/) ZENG Z.
 (KYAW/) KYAW H.
 (FISC/) FISCHER C L.
 (LIHH/) LI H.
 (SOPP/) SOPPET D R.
 (GENT/) GENTZ R L.
 (WEIY/) WEI Y.
 (MOOR/) MOORE P A.
 (YOIN/) YOUNG P E.
 (GREE/) GREENE J M.
 (FERR/) FERRIE A M.
 Ruben SM, Rosen CA, Li Y, Zeng Z, Kyaw H, Fischer CL, Li H;
 Soppet DR, Gentz RL, Wei Y, Moore PA, Young PE, Greene JM;
 Ferrie AM;
 WPI, 2003-310969/30.
 N-PSDB; ABX96975.
 New human secreted polypeptides and polynucleotides for diagnosing,
 prognosing, preventing and treating immune, hyperproliferative, liver,
 kidney, reproductive disorders and for identifying modulators of
 therapeutic use -
 Claim 11; Page 177; 209pp; English.
 The invention relates to an isolated polypeptide comprising an amino acid
 sequence at least 95% identical to sequence of 28 human secreted
 proteins, their fragments, polypeptide domain, epitope, secreted form,
 variant, allelic variant, or species homologue, or the encoded sequence
 included in ATCC 97921 and 97922. Also included are the encoding
 nucleic acids, recombinant vectors, host cells, antibodies, and genes.
 The proteins and nucleic acids are useful for diagnosing, preventing,
 treating, prognosing or ameliorating a medical condition e.g.
 immunodeficiencies (e.g. X-linked agammaglobulinemia, B cell
 immunodeficiencies), severe combined immunodeficiencies), autoimmune
 disorders (e.g. systemic erythematousus, rheumatoid arthritis, multiple

sclerosis, autoimmune thyroiditis, autoimmune haemolytic anaemia,
 Goodpasture's syndrome, Grave's disease, diabetes mellitus, dermatitis),
 haematopoietic disorders, inflammatory conditions (e.g. septic shock,
 sepsis, reperfusion injury, inflammatory bowel disease, Crohn's disease),
 respiratory disorders (e.g. asthma and allergy), gastrointestinal
 disorders, cancers (e.g. gastric, ovarian, lung, bladder, liver and
 breast), central nervous system (CNS) disorders (e.g. ischaemic brain
 injury and/or stroke, traumatic brain injury), neurodegenerative
 disorders (e.g. Parkinson's disease and Alzheimer's disease, AIDS-related
 dementia, and prion disease), cardiovascular disorders (e.g.
 atherosclerosis, myocarditis, cardiovascular disease, and cardiopulmonary
 bypass complications), inflammation (e.g. hepatitis, gout, trauma,
 pancreatitis, sarcoidosis, dermatitis, allogeneic transplant rejection),
 blood-related disorders (thrombosis, arterial thrombosis),
 hyperproliferative disorders, renal disorders (e.g. acute
 glomerulonephritis), endocrine disorders (e.g. Addison's disease,
 hyperthyroidism, hyperparathyroidism), liver diseases and disorders,
 reproductive system disorders (e.g. endometriosis), infectious diseases,
 and pancreatic disorders. Many other diseases and disorders are listed in
 the specification. They also useful as a vaccine adjuvant. Further they
 are useful to enhance or inhibit complement mediated cell lysis, for
 stimulating wound and tissue repair, angiogenesis, and the repair of
 vascular or lymphatic diseases or disorders. They are also useful
 to prevent hair loss, to modulate mammalian characteristics such as body
 height, weight, hair colour, and to increase or decrease storage
 capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 minerals, cofactors or other nutritional components. The proteins are
 also useful for identifying binding partners. The present sequence
 represents a secreted protein of the invention.

Sequence 100 AA;

Query Match 3.6%; Score 7; DB 24; Length 100;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 NKISLES 150
 DB 53 NKISLES 59

RESULT 15
 AAY65396
 ID AAY65396 standard; Protein; 101 AA.

XX AAY65396;

DT 01-FEB-2000 (first entry)

DE Human 5' EST related polypeptide SEQ ID NO:1557.

KW Human, 5' EST, expressed sequence tag, secreted protein, diagnosis;

KW gene therapy; chromosome mapping; upstream regulatory sequence;

KW forensic; location; development; protein synthesis; stability;

KW regulation; identification.

OS Homo sapiens.

PN WO9953051-A2.

PD 21-OCT-1999.

PF 09-APR-1999; 99WO-1B00712.

PR 09-APR-1998; 98US-0057719.

PR 28-APR-1998; 98US-0069047.

PA (GENT) GENSET.

PI Dumas Mline Edwards J, Duclert A, Giordano J;

DR WPI, 2000-038446/03.

DR N-PSDB; AA243010.

XX Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
PS Claim 3; Page 818-819; 837bp; English.
CC AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AAY64651 to
CC AAY65438 represent the EST-related proteins corresponding to AA242265 to
CC AA243052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AA242249 to AA242264 and AAY64644 to AAY64650 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 101 AA;
Query Match 3.6%; Score 7; DB 21; Length 101;
Best Local Similarity 100.0%; Pred.No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 144 NKISLES 150
DB 59 NKISLES 65
RESULT 16
AAY36137
ID AAY36137 standard; protein; 106 AA.
XX
AC AAY36137;
XX
DT 23-SEP-1999 (first entry)
XX
DE Human secreted protein #9.
XX
KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;
XX diagnostic; gene therapy; chromosome mapping; secretion vector.
XX
OS Homo sapiens.
XX
PN WO925825-A2.
XX
PD 27-MAY-1999.
XX
PF 13-NOV-1998; 98WO-IB01862.
XX
XX 04-SEP-1998; 98US-0099273.
XX 13-NOV-1997; 97US-0066677.
XX 17-DEC-1997; 97US-0069957.
XX 09-FEB-1998; 98US-0074121.
XX 13-APR-1998; 98US-0081563.
XX 10-AUG-1998; 98US-0096116.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX WPI, 1999-347472/29.
XX DR N-PSDB; AAX97821.
XX

PT Extended cDNAs encoding secreted proteins
XX
PS Example 28; Page 235-236; 307bp; English.
XX
CC AAY36129-Y36222 represent novel human secreted proteins encoded by the
CC extended cDNA sequences represented in AAX97813-X97906. The proteins
CC of the invention have cytosolic, thrombotic and osteopathic activity.
CC The extended cDNAs can be used to express secreted proteins or parts of
CC them or to obtain antibodies capable of binding to the secreted
CC proteins. They may also be used in diagnostic, forensic, gene therapy
CC and chromosome mapping procedures. Uses also include design of
CC expression vectors and secretion vectors.
XX
SQ Sequence 106 AA;
Query Match 3.6%; Score 7; DB 20; Length 106;
Best Local Similarity 100.0%; Pred.No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 144 NKISLES 150
DB 59 NKISLES 65
RESULT 17
AAY36184
ID AAY36184 standard; protein; 106 AA.
XX
AC AAY36184;
XX
DT 23-SEP-1999 (first entry)
XX
DE Human secreted protein #56.
XX
KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;
XX diagnostic; gene therapy; chromosome mapping; secretion vector.
XX
OS Homo sapiens.
XX
PN WO925825-A2.
XX
PD 27-MAY-1999.
XX
PF 13-NOV-1998; 98WO-IB01862.
XX
XX 04-SEP-1998; 98US-0099273.
XX 13-NOV-1997; 97US-0066677.
XX 17-DEC-1997; 97US-0069957.
XX 09-FEB-1998; 98US-0074121.
XX 13-APR-1998; 98US-0081563.
XX 10-AUG-1998; 98US-0096116.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX WPI, 1999-347472/29.
XX DR N-PSDB; AAX97868.
XX
XX Extended cDNAs encoding secreted proteins
XX
PS Claim 7; Page 290; 307bp; English.
XX
CC AAY36129-Y36222 represent novel human secreted proteins encoded by the
CC extended cDNA sequences represented in AAX97813-X97906. The proteins
CC of the invention have cytosolic, thrombotic and osteopathic activity.
CC The extended cDNAs can be used to express secreted proteins or parts of
CC them or to obtain antibodies capable of binding to the secreted
CC proteins. They may also be used in diagnostic, forensic, gene therapy
CC and chromosome mapping procedures. Uses also include design of
CC expression vectors and secretion vectors.
XX
SQ Sequence 106 AA;
XX

Query Match 3.6%; Score 7; DB 20; Length 106;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 NKISLES 150
 DB 59 NKISLES 65

RESULT 18

ID AAY35987 standard; Protein; 106 AA.

AC AAY35987;

DT 13-SEP-1999 (first entry)

DE Extended human secreted protein sequence, SEQ ID NO. 236.

XX Secreted protein; human; cytokine; cellular proliferation; cell movement;

XX cellular differentiation; immune system regulator; anti-inflammatory;

KM haematopoiesis regulator; tissue growth regulator; tumour inhibitor;

KM reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;

XX genetic disease.

XX Homo sapiens.

PN WO931236-A2.

PD 24-JUN-1999.

PF 17-DEC-1998; 98WO-1B02122.

PR 10-AUG-1998; 98US-0096116.

PR 17-DEC-1997; 97US-0069957.

PR 09-FEB-1998; 98US-0074121.

PR 13-APR-1998; 98US-0081563.

XX (GEST) GENSET.

PI Bougueleret L, Duclet A, Dumas Milne Edwards J;

DR WPI; 1999-385906/32.

DR N-PSDB; AAY97671.

PT New isolated human secreted proteins

XX Claim 9; Page 264; 516bp; English.

XX This sequence is encoded by an extended human secreted protein coding
 CC sequence of the invention. The secreted proteins can be used in treating
 CC or controlling a variety of human conditions. The secreted proteins may
 CC act as cytokines or may affect cellular proliferation or differentiation
 CC or may act as immune system regulators, haematopoiesis regulators, tissue
 CC growth regulators, regulators of reproductive hormones or cell movement
 CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
 CC tumour inhibition activity. The DNAs can be used in forensic procedures
 CC to identify individuals or in diagnostic procedures to identify
 CC individuals having genetic diseases resulting from abnormal expression of
 CC the genes corresponding to the extended cDNAs. They are also useful for
 CC constructing a high resolution map of the human chromosomes. They can
 CC also be used for gene therapy to control or treat genetic diseases.

SO Sequence 106 AA;

Query Match 3.6%; Score 7; DB 20; Length 106;

Best Local Similarity 100.0%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 NKISLES 150
 DB 59 NKISLES 65

RESULT 19

ID ABB97399 standard; Protein; 106 AA.

AC ABB97399;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 667.

XX Human; antianemic; vulnery; antiinflammatory; immunomodulator;

KM antiinfectivity; cerebroprotective; cytostatic; rheumatic; gene therapy;

KM neuroprotective; antiparkinsonian; protein therapy; ESR;

XX expressed sequence tag.

XX Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI; 2002-292408/33.

DR N-PSDB; ABN32585.

PT An isolated polynucleotide for treating diseases associated with its

XX encoded polypeptide such as cancer and multiple sclerosis -

XX Claim 20; SEQ ID NO 667; 509bp; English.

XX The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention.

SO Sequence 106 AA;

Query Match 3.6%; Score 7; DB 23; Length 106;

Best Local Similarity 100.0%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 NKISLES 150
 DB 59 NKISLES 65

RESULT 20

ID ABG60021 standard; Protein; 115 AA.

AC ABG60021;

DT 30-JUL-2002 (first entry)

DE Human DITHP polypeptide #79.

XX Human; DITHP; diagnostic and therapeutic polypeptide; bone; testis; skin;

KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
 KW inflammatory disorder; viral infection; bacterial infection; seizure;
 KW fungal infection; parasitic infections; developmental disorder; breast;
 KW endocrine disorder; metabolic disorder; neurological disorder; cervix;
 KW gastrointestinal disorder; transport disorder; gene therapy; kidney;
 KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
 KW thymus.
 XX
 OS Homo sapiens.
 XX
 PN WO200220754-A2.
 XX
 PD 14-MAR-2002.
 XX
 PE 29-AUG-2001; 2001WO-US7127.
 XX
 PR 05-SEP-2000; 2000US-229747P.
 PR 05-SEP-2000; 2000US-229748P.
 PR 05-SEP-2000; 2000US-229749P.
 PR 05-SEP-2000; 2000US-229750P.
 PR 05-SEP-2000; 2000US-229751P.
 PR 05-SEP-2000; 2000US-230583P.
 PR 06-SEP-2000; 2000US-230585P.
 PR 06-SEP-2000; 2000US-230514P.
 PR 06-SEP-2000; 2000US-230515P.
 PR 06-SEP-2000; 2000US-230517P.
 PR 06-SEP-2000; 2000US-230518P.
 PR 06-SEP-2000; 2000US-230519P.
 PR 06-SEP-2000; 2000US-230595P.
 PR 06-SEP-2000; 2000US-230597P.
 PR 06-SEP-2000; 2000US-230598P.
 PR 06-SEP-2000; 2000US-230599P.
 PR 06-SEP-2000; 2000US-230610P.
 PR 06-SEP-2000; 2000US-230656P.
 PR 06-SEP-2000; 2000US-230988P.
 PR 07-SEP-2000; 2000US-230951P.
 PR 07-SEP-2000; 2000US-231163P.
 PR 07-SEP-2000; 2000US-231167P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup NS, Hillman JL,
 PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
 PI Momiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
 PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
 PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
 XX
 DR WPI; 2002-383054/41.
 DR N-P-SDB; ABK71612.
 XX
 PT An isolated polynucleotide useful in diagnostics and therapeutics -
 XX
 PS Claim 29, Page 578; 686bp; English.
 XX
 CC The invention relates to human diagnostic and therapeutic (dithp)
 CC polynucleotides and their associated polypeptides (DITHP polypeptides).
 CC The sequences of the invention are used in the treatment and diagnosis of
 CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
 CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,
 CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
 CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
 CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal
 CC infections, parasitic infections, developmental disorders (e.g. anaemia,
 CC epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
 CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders
 CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,
 CC amyotrophic lateral sclerosis), multiple sclerosis), gastrointestinal
 CC disorders (e.g. ulcerative colitis, lysinuria) and transport disorders
 CC (e.g. myotonic dystrophy, catatonia, peripheral neuropathy). Sequences
 CC ABG5943;ABG60220 represent human DITHP polypeptides of the invention.
 XX
 XX Sequence 115 AA;

Query Match 3.6%; Score 7; DB 23; Length 115;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 81 TKTCFEK 87
 Db 66 TKTCFEK 72
 XX
 RESULT 21
 ABR41138
 ID ABR41138 standard; Protein; 128 AA.
 XX
 AC ABR41138;
 XX
 DT 02-JUN-2003 (first entry)
 XX
 DE Human DITHP enzyme.
 XX
 KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging;
 KW enzyme.
 XX
 OS Homo sapiens.
 XX
 PN WO200297031-A2.
 XX
 PD 05-DEC-2002.
 XX
 PE 27-MAR-2002; 2002WO-US10056.
 XX
 PR 28-MAR-2001; 2001US-279619P.
 PR 29-MAR-2001; 2001US-280067P.
 PR 29-MAR-2001; 2001US-280068P.
 PR 16-MAY-2001; 2001US-291280P.
 PR 17-MAY-2001; 2001US-291839P.
 PR 17-MAY-2001; 2001US-291849P.
 PR 19-JUN-2001; 2001US-299428P.
 PR 20-JUN-2001; 2001US-299776P.
 PR 20-JUN-2001; 2001US-300001P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J,
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Anshey SR;
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B,
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX
 DR WPI; 2003-129518/12.
 DR N-P-SDB; ACC46082.
 XX
 PT Novel human diagnostic and therapeutic polypeptide useful for
 PT identifying test compound which specifically binds to a polypeptide
 PT encoded by human diagnostic and therapeutic polynucleotide; and to
 PT induce antibodies -
 XX
 PS Claim 27, SEQ ID No 673; 591bp; English.
 XX
 CC The invention relates to novel human diagnostic and therapeutic
 CC polynucleotides designated dithp (ACC46080-ACC46749) and to their
 CC encoded proteins (DITHP; ABR41136-ABR41812). The invention also relates
 CC to polynucleotide sequences at least 90% identical to the dithp cDNA
 CC sequences of the invention; recombinant vectors, host cells and
 CC transgenic organisms comprising a dithp nucleic acid sequence; the
 CC recombinant production of DITHP proteins; antibodies specific for DITHP
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods
 CC of detecting dithp nucleotide and protein sequences; methods of screening

CC	The present invention relates to the genome sequence of <i>Listeria</i>
CC	monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC	it are useful for selecting probes and primers for detecting genes in <i>L.</i>
CC	monocytogenes and related organisms, and for studying genetic
CC	polymorphisms and other genomes. The present sequence is a protein
CC	encoded by the genome sequence of the present invention. Proteins
CC	expressed from the genome sequence are useful for raising specific
CC	antibodies, identification of <i>L. monocytogenes</i> and related organisms, and
CC	for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC	B12. The genome sequence and proteins encoded by it are also useful for
CC	selecting compounds that regulate gene expression and cell replication
CC	and modulate <i>L. monocytogenes</i> -related diseases. In addition, the genome
CC	sequence and proteins encoded by it are useful in pharmaceutical and
CC	vaccines compositions for the treatment or prevention of infections by <i>L.</i>
CC	monocytogenes and related organisms.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIP0
CC	at ftp.wipo.int/pub/published_pct_sequences.
CC	
SO	Sequence 140 AA;
SO	
Query Match	3.6%; Score 7; DB 23; Length 140;
Best Local Similarity	100.0%; Pred. No. 67;
Matches 7; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	33 ILKNIST 39
Db	131 ILKNIST 137
RESULT 24	
ABG63138	ABG63138 standard; Protein; 146 AA.
ID	
AC	ABG63138;
XX	
DT	22-AUG-2002 (first entry)
XX	
DE	<i>Bacillus</i> glyphosate N-acetyl transferase #183.
XX	
KW	Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KW	N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
XX	transgenic; weed control; cotton; corn; soybean.
XX	
OS	<i>Bacillus</i> sp.
XX	
PN	WO200236782-A2.
XX	
PD	10-MAY-2002.
XX	
PF	29-OCT-2001; 2001WO-US46227.
XX	
PR	30-OCT-2000; 2000US-244385P.
XX	
PA	(MAXY-) MAXYGEN INC.
PA	(PTON-) PTONEER HI-BRED INT INC.
XX	
PI	Castle LA, Siehl D, Giver LD, Minshall J, Ivy C, Chen YH;
PI	Duck NB;
XX	
XX	WPI; 2002-490010/52.
DR	N-PSDB; ABR92884.
XX	
PT	New glyphosate-N-acetyltransferase gene and polypeptide having
PT	increased rate of catalysis and increased stability, useful for
PT	generating glyphosate resistant plants
XX	
XX	Claim 127; Page 191; 254pp; English.
XX	
CC	The invention relates to an isolated or recombinant polypeptide (I)
CC	comprising an amino acid sequence that can be optimally aligned with
CC	a sequence (S) of 147 amino acids fully defined in the specification
CC	to generate a similarity score of at least 430, using the BLOSUM62

CC	matrix,a gap existence penalty of 11, and a gap extension penalty
CC	of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC	activity, including fragments, variants, polypeptides recognised by (I)
CC	anti-(S) antibodies. Also included are the polynucleotides encoding (I)
CC	(including variant and recombinant polynucleotides)', a cell comprising
CC	the recombinant polynucleotide heterologous to the cell, a transgenic
CC	plant or its explant and seed comprising the above cell and producing (M)
CC	a variant of (I) by recursively recombining (I) with a second
CC	polynucleotide, forming a library of variant polynucleotides.
CC	The recombinant GAT polynucleotide is useful for producing a glyphosate
CC	(N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC	These transgenic plants are useful for selectively controlling weeds in
CC	field containing a crop, and for producing crops such as cotton, corn or
CC	soybean. The transgenic plant is useful for controlling weeds in a field
CC	containing a crop and for preventing emergence of herbicide resistant
CC	weeds in a field. The present sequence represents a GAT polypeptide
CC	of the invention.
CC	
XX	
SQ	Sequence 146 AA;
QY	
Query Match	3.6%; Score 7; DB 23; Length 146;
Best Local Similarity	100.0%; Pred. No. 70;
Matches 7; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
OY	170 SSLIKHA 176
Db	88 SSLIKHA 94
RESULT 25	
ABG63139	ID ABG63139 standard; Protein; 146 AA.
XX	
XX	ABG63139;
XX	
DT	22-AUG-2002 (first entry)
XX	
DE	Bacillus glyphosate N-acetyl transferase #184.
XX	
KX	Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KW	N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
KW	transgenic; weed control; cotton; corn; soybean.
XX	
OS	Bacillus sp.
PX	
PN	WO200236782-A2.
XX	
PD	10-MAY-2002.
PF	29-OCT-2001; 2001MO-US46227.
XX	
PR	30-OCT-2000; 2000US-244385P.
XX	
PPA	(MAXY-) MAXYGEN INC.
XX	(PION-) PIONEER HI-BRED INT INC.
PI	Caele LA, Siehl D, Gliver LT, Minshull J, Ivy C, Chen YH;
PI	Duck NB;
XX	
DR	WPI; 2002-490010/52.
XX	
XX	N-P8DB; ABR92885.
PT	New glyphosate-N-acetyltransferase gene and polypeptide having
PT	increased rate of catalysis and increased stability, useful for
PT	generating glyphosate resistant plants -
XX	
PS	Claim 127; Page 191; 254pp; English.
XX	
CC	The invention relates to an isolated or recombinant polypeptide (I)
CC	comprising an amino acid sequence that can be optimally aligned with
CC	a sequence (S) of 147 amino acids fully defined in the specification
CC	to generate a similarity score of at least 430, using the BLOSUM62
CC	matrix, a gap existence penalty of 11, and a gap extension penalty

CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (I) by recursively recombining (I) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
XX
XX
SQ Sequence 146 AA;
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLLKHA 176
DB 88 SSLLKHA 94
RESULT 26
ABG63143
ID ABG63143 standard; Protein; 146 AA.
XX
XX ABG63143;
AC
XX
XX 22-AUG-2002 (first entry)
DT
XX
XX Bacillus glyphosate N-acetyl transferase #188.
DE
XX
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
KW transgenic; weed control; cotton; corn; soybean.
XX
XX
OS Bacillus sp.
XX
XX MO200236782-A2.
PN
XX
XX 10-MAY-2002.
PD
XX
XX 29-OCT-2001; 2001MO-US46227.
PF
XX
XX 30-OCT-2000; 2000US-244385P.
PR
XX
XX (MAXY-) MAXYGEN INC.
PA (PION-) PIONEER HI-BRED INT INC.
XX
XX Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
PI Duck NB;
XX
XX WPI; 2002-490010/52.
DR N-PSDB; ABK92889.
XX
XX
PT New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants -
XX
XX
PS Claim 127; Page 191; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (I)
XX comprising an amino acid sequence that can be optimally aligned with
XX a sequence (S) of 147 amino acids fully defined in the specification
XX to generate a similarity score of at least 430, using the BLOSUM62
XX matrix, a gap existence penalty of 11, and a gap extension penalty
XX of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)

CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (I) by recursively recombining (I) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
XX
XX
SQ Sequence 146 AA;
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLLKHA 176
DB 88 SSLLKHA 94
RESULT 27
ABG63150
ID ABG63150 standard; Protein; 146 AA.
XX
XX ABG63150;
AC
XX
XX 22-AUG-2002 (first entry)
DT
XX
XX Bacillus glyphosate N-acetyl transferase #195.
DE
XX
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
KW transgenic; weed control; cotton; corn; soybean.
XX
XX
OS Bacillus sp.
XX
XX MO200236782-A2.
PN
XX
XX 10-MAY-2002.
PD
XX
XX 29-OCT-2001; 2001MO-US46227.
PF
XX
XX 30-OCT-2000; 2000US-244385P.
PR
XX
XX (MAXY-) MAXYGEN INC.
PA (PION-) PIONEER HI-BRED INT INC.
XX
XX Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
PI Duck NB;
XX
XX WPI; 2002-490010/52.
DR N-PSDB; ABK92896.
XX
XX
PT New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants -
XX
XX
PS Claim 127; Page 192; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (I)
XX comprising an amino acid sequence that can be optimally aligned with
XX a sequence (S) of 147 amino acids fully defined in the specification
XX to generate a similarity score of at least 430, using the BLOSUM62
XX matrix, a gap existence penalty of 11, and a gap extension penalty
XX of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
XX activity, including fragments, variants, polypeptides recognised by

CC anti-(S) antibodies. Also included are the polynucleotides encoding (1)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (1) by recursively recombining (1) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
CC
XX
SQ Sequence 146 AA;
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLIKHA 176
DB 88 SSLIKHA 94
RESULT 28
ABG63155
ID ABG63155 standard; Protein; 146 AA.
AC ABG63155;
XX
XX 22-AUG-2002 (first entry)
DT
XX
DE Bacillus glyphosate N-acetyl transferase #200.
XX
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
XX N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
XX transgenic; weed control; cotton; corn; soybean.
XX
XX Bacillus sp.
XX
XX WO200236782-A2.
XX
XX 10-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-US46227.
XX
XX 30-OCT-2000; 2000US-244385P.
XX
XX (MAXY-) MAXYGEN INC.
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
XX Duck NB;
XX
XX WPI; 2002-490010/52.
XX N-PSDB; ABK92901.
XX
XX New glyphosate-N-acetyltransferase gene and polypeptide having
XX increased rate of catalysis and increased stability, useful for
XX generating glyphosate resistant plants -
XX
XX Claim 127; Page 192; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (1)
XX comprising an amino acid sequence that can be optimally aligned with
XX a sequence (S) of 147 amino acids fully defined in the specification
XX to generate a similarity score of at least 430, using the BLOSUM62
XX matrix, a gap existence penalty of 11, and a gap extension penalty
XX of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
XX activity, including fragments, variants, polypeptides recognised by
XX anti-(S) antibodies. Also included are the polynucleotides encoding (1)

CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (1) by recursively recombining (1) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
CC
XX
SQ Sequence 146 AA;
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLIKHA 176
DB 88 SSLIKHA 94
RESULT 29
ABG63156
ID ABG63156 standard; Protein; 146 AA.
AC ABG63156;
XX
XX 22-AUG-2002 (first entry)
DT
XX
DE Bacillus glyphosate N-acetyl transferase #201.
XX
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
XX N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
XX transgenic; weed control; cotton; corn; soybean.
XX
XX Bacillus sp.
XX
XX WO200236782-A2.
XX
XX 10-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-US46227.
XX
XX 30-OCT-2000; 2000US-244385P.
XX
XX (MAXY-) MAXYGEN INC.
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
XX Duck NB;
XX
XX WPI; 2002-490010/52.
XX N-PSDB; ABK92902.
XX
XX New glyphosate-N-acetyltransferase gene and polypeptide having
XX increased rate of catalysis and increased stability, useful for
XX generating glyphosate resistant plants -
XX
XX Claim 127; Page 192; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (1)
XX comprising an amino acid sequence that can be optimally aligned with
XX a sequence (S) of 147 amino acids fully defined in the specification
XX to generate a similarity score of at least 430, using the BLOSUM62
XX matrix, a gap existence penalty of 11, and a gap extension penalty
XX of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
XX activity, including fragments, variants, polypeptides recognised by
XX anti-(S) antibodies. Also included are the polynucleotides encoding (1)
XX (including variants and recombinant polynucleotides), a cell comprising

CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (I) by recursively recombining (I) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
CC
SQ Sequence 146 AA;
XX
XX
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLIKHA 176
|||
|||
DB 88 SSLIKHA 94
RESULT 30
ABG63160
ID ABG63160 standard; Protein; 146 AA.
XX
XX ABG63160;
XX
XX 22-AUG-2002 (first entry)
XX
XX Bacillus glyphosate N-acetyl transferase #205.
XX
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
XX N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
XX transgenic; weed control; cotton; corn; soybean.
XX
XX Bacillus sp.
XX
XX WO200236782-A2.
XX
XX 10-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-US46227.
XX
XX 30-OCT-2000; 2000US-244385P.
XX
XX (MAXY-) MAXYGEN INC.
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Castle LA, Siehl D, Glver LJ, Minshull J, Ivy C, Chen YH;
XX Duck NB;
XX
XX MPI; 2002-490010/52.
XX N-PSDB; ABK92906.
XX
XX New glyphosate-N-acetyltransferase gene and polypeptide having
XX increased rate of catalysts and increased stability, useful for
XX generating glyphosate resistant plants
XX
XX Claim 127; Page 193; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (I)
XX comprising an amino acid sequence that can be optimally aligned with
XX a sequence (S) of 147 amino acids fully defined in the specification
XX to generate a similarity score of at least 430, using the BLOSUM62
XX matrix, a gap existence penalty of 11, and a gap extension penalty
XX of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
XX activity, including fragments, variants, polypeptides recognised by
XX anti-(S) antibodies. Also included are the polynucleotides encoding (I)
XX (including variants and recombinant polynucleotides), a cell comprising
XX the recombinant polynucleotide heterologous to the cell, a transgenic

CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (I) by recursively recombining (I) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
CC
SQ Sequence 146 AA;
XX
XX
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLIKHA 176
|||
|||
DB 88 SSLIKHA 94
RESULT 31
ABG63162
ID ABG63162 standard; Protein; 146 AA.
XX
XX ABG63162;
XX
XX 22-AUG-2002 (first entry)
XX
XX Bacillus glyphosate N-acetyl transferase #207.
XX
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
XX N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
XX transgenic; weed control; cotton; corn; soybean.
XX
XX Bacillus sp.
XX
XX WO200236782-A2.
XX
XX 10-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-US46227.
XX
XX 30-OCT-2000; 2000US-244385P.
XX
XX (MAXY-) MAXYGEN INC.
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Castle LA, Siehl D, Glver LJ, Minshull J, Ivy C, Chen YH;
XX Duck NB;
XX
XX MPI; 2002-490010/52.
XX N-PSDB; ABK92908.
XX
XX New glyphosate-N-acetyltransferase gene and polypeptide having
XX increased rate of catalysts and increased stability, useful for
XX generating glyphosate resistant plants
XX
XX Claim 127; Page 193; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (I)
XX comprising an amino acid sequence that can be optimally aligned with
XX a sequence (S) of 147 amino acids fully defined in the specification
XX to generate a similarity score of at least 430, using the BLOSUM62
XX matrix, a gap existence penalty of 11, and a gap extension penalty
XX of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
XX activity, including fragments, variants, polypeptides recognised by
XX anti-(S) antibodies. Also included are the polynucleotides encoding (I)
XX (including variants and recombinant polynucleotides), a cell comprising
XX the recombinant polynucleotide heterologous to the cell, a transgenic
XX plant or its explant and seed comprising the above cell and producing (M)

CC a variant of (1) by recursively recombining (1) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
XX
SQ Sequence 146 AA;
XX
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 170 SSLIKHA 176
DB 88 SSLIKHA 94
|||||
|
RESULT 32
ABG63163
ID ABG63163 standard; Protein; 146 AA.
XX
AC ABG63163;
XX
DT 22-AUG-2002 (first entry)
XX
DE Bacillus glyphosate N-acetyl transferase #208.
XX
KM Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
XX N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
XX transgenic; weed control; cotton; corn; soybean.
XX
OS Bacillus sp.
XX
PN WO200236782-A2.
XX
PD 10-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-US46227.
XX
PR 30-OCT-2000; 2000US-244385P.
XX
PA (MAXY-) MAXYGEN INC.
XX (PION-) PIONEER HI-BRED INT INC.
XX
PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
XX Duck NB;
XX
DR WPI; 2002-490010/52.
XX N-PSDB; ABK92909.
XX
XX New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants
XX
PS Claim 127; Page 193; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (1)
CC comprising an amino acid sequence that can be optimally aligned with
CC a sequence (S) of 147 amino acids fully defined in the specification
CC to generate a similarity score of at least 430, using the BLOSUM62
CC matrix, a gap existence penalty of 11, and a gap extension penalty
CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (1)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing
CC a variant of (1) by recursively recombining (1) with a second

CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
XX
SQ Sequence 146 AA;
XX
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 170 SSLIKHA 176
DB 88 SSLIKHA 94
|||||
|
RESULT 33
ABG63168
ID ABG63168 standard; Protein; 146 AA.
XX
AC ABG63168;
XX
DT 22-AUG-2002 (first entry)
XX
DE Bacillus glyphosate N-acetyl transferase #213.
XX
KM Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
XX N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
XX transgenic; weed control; cotton; corn; soybean.
XX
OS Bacillus sp.
XX
PN WO200236782-A2.
XX
PD 10-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-US46227.
XX
PR 30-OCT-2000; 2000US-244385P.
XX
PA (MAXY-) MAXYGEN INC.
XX (PION-) PIONEER HI-BRED INT INC.
XX
PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
XX Duck NB;
XX
DR WPI; 2002-490010/52.
XX N-PSDB; ABK92914.
XX
XX New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants
XX
PS Claim 127; Page 193; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (1)
CC comprising an amino acid sequence that can be optimally aligned with
CC a sequence (S) of 147 amino acids fully defined in the specification
CC to generate a similarity score of at least 430, using the BLOSUM62
CC matrix, a gap existence penalty of 11, and a gap extension penalty
CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (1)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing
CC a variant of (1) by recursively recombining (1) with a second
CC polynucleotide, forming a library of variant polynucleotides.

CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
CC
SQ Sequence 146 AA:
XX
XX
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 170 SSLIKHA 176
Db 88 SSLIKHA 94
|||||
RESULT 34
ABG63171
ID ABG63171 standard; Protein; 146 AA.
XX
AC ABG63171;
XX
DT 22-AUG-2002 (first entry)
XX
DE Bacillus glyphosate N-acetyl transferase #216.
XX
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
KM transgenic; weed control; cotton; corn; soybean.
XX
OS Bacillus sp.
XX
PN WO200236782-A2.
XX
PD 10-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-US46227.
XX
PR 30-OCT-2000; 2000US-244385P.
XX
PA (MAXY-) MAXYGEN INC.
PI (PION-) PIONEER HI-BRED INT INC.
XX
PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
PI Duck NB;
XX
XX WPI; 2002-490010/52.
DR N-PSDB; ABK92917.
XX
XX New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants -
PT
PS Claim 127; Page 193-194; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (I)
CC comprising an amino acid sequence that can be optimally aligned with
CC a sequence (S) of 147 amino acids fully defined in the specification
CC to generate a similarity score of at least 430, using the BLOSUM62
CC matrix, a gap existence penalty of 11, and a gap extension penalty
CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (I) by recursively recombining (I) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate

CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
CC
SQ Sequence 146 AA:
XX
XX
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 170 SSLIKHA 176
Db 88 SSLIKHA 94
|||||
RESULT 35
ABG63183
ID ABG63183 standard; Protein; 146 AA.
XX
AC ABG63183;
XX
DT 22-AUG-2002 (first entry)
XX
DE Bacillus glyphosate N-acetyl transferase #228.
XX
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
KM transgenic; weed control; cotton; corn; soybean.
XX
OS Bacillus sp.
XX
PN WO200236782-A2.
XX
PD 10-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-US46227.
XX
PR 30-OCT-2000; 2000US-244385P.
XX
PA (MAXY-) MAXYGEN INC.
PI (PION-) PIONEER HI-BRED INT INC.
XX
PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
PI Duck NB;
XX
XX WPI; 2002-490010/52.
DR N-PSDB; ABK92929.
XX
XX New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants -
PT
PS Claim 127; Page 194; 254pp; English.
XX
XX The invention relates to an isolated or recombinant polypeptide (I)
CC comprising an amino acid sequence that can be optimally aligned with
CC a sequence (S) of 147 amino acids fully defined in the specification
CC to generate a similarity score of at least 430, using the BLOSUM62
CC matrix, a gap existence penalty of 11, and a gap extension penalty
CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (I) by recursively recombining (I) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.

CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.

XX SQ Sequence 146 AA;

Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||
88 SSLIKHA 94

Db 88 SSLIKHA 94

RESULT 36

ABG63185 ID ABG63185 standard; Protein; 146 AA.

XX AC ABG63185;

DT 22-AUG-2002 (first entry)

DE Bacillus glyphosate N-acetyl transferase #230.

XX KW Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
XX KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
XX transgenic; weed control; cotton; corn; soybean.

OS Bacillus sp.

XX PN W0200236782-A2.

XX PD 10-MAY-2002.

PF 29-OCT-2001; 2001WO-US46227.

XX PR 30-OCT-2000; 2000US-244385P.

XX PA (MAXY-) MAXYGEN INC.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;

XX PI Duck NB;

XX DR WPI; 2002-490010/52.

XX DR N-PSDB; ABK92931.

PT New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants

XX PS Claim 127; Page 195; 254pp; English.

XX CC The invention relates to an isolated or recombinant polypeptide (I)
CC comprising an amino acid sequence that can be optimally aligned with
CC a sequence (S) of 147 amino acids fully defined in the specification
CC to generate a similarity score of at least 430, using the BLOSUM62
CC matrix, a gap existence penalty of 11, and a gap extension penalty
CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (I) by recursively recombining (I) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in

CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.

XX SQ Sequence 146 AA;

Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||
88 SSLIKHA 94

Db 88 SSLIKHA 94

RESULT 37

ABG63186 ID ABG63186 standard; Protein; 146 AA.

XX AC ABG63186;

DT 22-AUG-2002 (first entry)

DE Bacillus glyphosate N-acetyl transferase #231.

XX KW Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
XX KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
XX transgenic; weed control; cotton; corn; soybean.

OS Bacillus sp.

XX PN W0200236782-A2.

XX PD 10-MAY-2002.

PF 29-OCT-2001; 2001WO-US46227.

XX PR 30-OCT-2000; 2000US-244385P.

XX PA (MAXY-) MAXYGEN INC.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;

XX PI Duck NB;

XX DR WPI; 2002-490010/52.

XX DR N-PSDB; ABK92932.

PT New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants

XX PS Claim 127; Page 195; 254pp; English.

XX CC The invention relates to an isolated or recombinant polypeptide (I)
CC comprising an amino acid sequence that can be optimally aligned with
CC a sequence (S) of 147 amino acids fully defined in the specification
CC to generate a similarity score of at least 430, using the BLOSUM62
CC matrix, a gap existence penalty of 11, and a gap extension penalty
CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (I) by recursively recombining (I) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or

CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.

XX Sequence 146 AA;

SO Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||||

DB 88 SSLIKHA 94

RESULT 38

ABG63187
ID ABG63187 standard; Protein; 146 AA.

XX ABG63187;

AC 22-AUG-2002 (first entry)

DT Bacillus glyphosate N-acetyl transferase #232.

DE Glyphosate resistance; glyphosate N-acetyltransferase; GAT;

XX N-phosphonomethylglycine; herbicide resistance; plant; enzyme;

KW transgenic; weed control; cotton; corn; soybean.

XX Bacillus sp.

OS WO200236782-A2.

PN 10-MAY-2002.

PD 29-OCT-2001; 2001WO-US46227.

PF 30-OCT-2000; 2000US-244385P.

PR (MAXY-) MAXYGEN INC.

PA (PION-) PIONEER HI-BRED INT INC.

XX Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;

PI Duck NB;

DR WPI; 2002-490010/52.

DR N-PSDB; ABK92933.

PT New glyphosate-N-acetyltransferase gene and polypeptide having

PT increased rate of catalysis and increased stability, useful for

PT generating glyphosate resistant plants

XX Claim 127; Page 195; 254pp; English.

XX The invention relates to an isolated or recombinant polypeptide (I)

CC comprising an amino acid sequence that can be optimally aligned with

CC a sequence (S) of 147 amino acids fully defined in the specification

CC to generate a similarity score of at least 430, using the BLOSUM62

CC matrix, a gap existence penalty of 11, and a gap extension penalty

CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)

CC activity, including fragments, variants, polypeptides recognised by

CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)

CC (including variants and recombinant polynucleotides), a cell comprising

CC the recombinant polynucleotide heterologous to the cell, a transgenic

CC plant or its explant and seed comprising the above cell and producing (M)

CC a variant of (I) by recursively recombining (I) with a second

CC polynucleotide, forming a library of variant polynucleotides.

CC The recombinant GAT polynucleotide is useful for producing a glyphosate

CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.

CC These transgenic plants are useful for selectively controlling weeds in

CC field containing a crop, and for producing crops such as cotton, corn or

CC soybean. The transgenic plant is useful for controlling weeds in a field

CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.

XX Sequence 146 AA;

SO Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||||

DB 88 SSLIKHA 94

RESULT 39

ABG63188
ID ABG63188 standard; Protein; 146 AA.

XX ABG63188;

AC 22-AUG-2002 (first entry)

DT Bacillus glyphosate N-acetyl transferase #233.

DE Glyphosate resistance; glyphosate N-acetyltransferase; GAT;

XX N-phosphonomethylglycine; herbicide resistance; plant; enzyme;

KW transgenic; weed control; cotton; corn; soybean.

XX Bacillus sp.

OS WO200236782-A2.

PN 10-MAY-2002.

PD 29-OCT-2001; 2001WO-US46227.

PF 30-OCT-2000; 2000US-244385P.

PR (MAXY-) MAXYGEN INC.

PA (PION-) PIONEER HI-BRED INT INC.

XX Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;

PI Duck NB;

DR WPI; 2002-490010/52.

DR N-PSDB; ABK92934.

PT New glyphosate-N-acetyltransferase gene and polypeptide having

PT increased rate of catalysis and increased stability, useful for

PT generating glyphosate resistant plants

XX Claim 127; Page 195; 254pp; English.

XX The invention relates to an isolated or recombinant polypeptide (I)

CC comprising an amino acid sequence that can be optimally aligned with

CC a sequence (S) of 147 amino acids fully defined in the specification

CC to generate a similarity score of at least 430, using the BLOSUM62

CC matrix, a gap existence penalty of 11, and a gap extension penalty

CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)

CC activity, including fragments, variants, polypeptides recognised by

CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)

CC (including variants and recombinant polynucleotides), a cell comprising

CC the recombinant polynucleotide heterologous to the cell, a transgenic

CC plant or its explant and seed comprising the above cell and producing (M)

CC a variant of (I) by recursively recombining (I) with a second

CC polynucleotide, forming a library of variant polynucleotides.

CC The recombinant GAT polynucleotide is useful for producing a glyphosate

CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.

CC These transgenic plants are useful for selectively controlling weeds in

CC field containing a crop, and for producing crops such as cotton, corn or

CC soybean. The transgenic plant is useful for controlling weeds in a field

CC weeds in a field. The present sequence represents a GAT polypeptide
CC of the invention.
XX
SQ Sequence 146 AA;
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLIKHA 176
DB 88 SSLIKHA 94
RESULT 40
ABG63189
ID ABG63189 standard; Protein; 146 AA.
AC ABG63189;
XX
DT 22-AUG-2002 (first entry)
XX
DE Bacillus glyphosate N-acetyl transferase #234.
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
KM transgenic; weed control; cotton; corn; soybean.
XX
OS Bacillus sp.
XX
PN WO200236782-A2.
XX
PD 10-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-US46227.
XX
PR 30-OCT-2000; 2000US-244385P.
XX
PA (MAXY-) MAXYGEN INC.
PI (PION-) PIONEER HI-BRED INT INC.
PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
PI Duck NB;
XX
DR WPI: 2002-490010/52.
DR N-PSDB; ABK92935.
PT New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants -
XX
PS Claim 127; Page 195; 254pp; English.
XX
CC The invention relates to an isolated or recombinant polypeptide (1)
CC comprising an amino acid sequence that can be optimally aligned with
CC a sequence (S) of 147 amino acids fully defined in the specification
CC to generate a similarity score of at least 430, using the BIOSIM62
CC matrix, a gap existence penalty of 11, and a gap extension penalty
CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (1)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (1) by recursively recombining (1) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide

CC of the invention.
XX
SQ Sequence 146 AA;
Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLIKHA 176
DB 88 SSLIKHA 94
RESULT 41
ABG63190
ID ABG63190 standard; Protein; 146 AA.
AC ABG63190;
XX
DT 22-AUG-2002 (first entry)
XX
DE Bacillus glyphosate N-acetyl transferase #235.
XX Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
KM transgenic; weed control; cotton; corn; soybean.
XX
OS Bacillus sp.
XX
PN WO200236782-A2.
XX
PD 10-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-US46227.
XX
PR 30-OCT-2000; 2000US-244385P.
XX
PA (MAXY-) MAXYGEN INC.
PI (PION-) PIONEER HI-BRED INT INC.
PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
PI Duck NB;
XX
DR WPI: 2002-490010/52.
DR N-PSDB; ABK92936.
PT New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants -
XX
PS Claim 127; Page 195; 254pp; English.
XX
CC The invention relates to an isolated or recombinant polypeptide (1)
CC comprising an amino acid sequence that can be optimally aligned with
CC a sequence (S) of 147 amino acids fully defined in the specification.
CC to generate a similarity score of at least 430, using the BIOSIM62
CC matrix, a gap existence penalty of 11, and a gap extension penalty
CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
CC activity, including fragments, variants, polypeptides recognised by
CC anti-(S) antibodies. Also included are the polynucleotides encoding (1)
CC (including variants and recombinant polynucleotides), a cell comprising
CC the recombinant polynucleotide heterologous to the cell, a transgenic
CC plant or its explant and seed comprising the above cell and producing (M)
CC a variant of (1) by recursively recombining (1) with a second
CC polynucleotide, forming a library of variant polynucleotides.
CC The recombinant GAT polynucleotide is useful for producing a glyphosate
CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
CC These transgenic plants are useful for selectively controlling weeds in
CC field containing a crop, and for producing crops such as cotton, corn or
CC soybean. The transgenic plant is useful for controlling weeds in a field
CC containing a crop and for preventing emergence of herbicide resistant
CC weeds in a field. The present sequence represents a GAT polypeptide

XX SQ Sequence 146 AA;
 Query Match 3.6%; Score 7; DB 23; Length 146;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 170 SSLIKHA 176
 |||||
 88 SSLIKHA 94
 DB

RESULT 42
 ABG63193
 ID ABG63193 standard; Protein; 146 AA.
 XX
 AC ABG63193;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Bacillus glyphosate N-acetyl transferase #238.
 XX
 KW Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
 KW N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
 KW transgenic; weed control; cotton; corn; soybean.
 OS
 XX Bacillus sp.
 XX
 PN MO200236782-A2.
 XX
 PD 10-MAY-2002.
 XX
 PF 29-OCT-2001; 2001MO-US46227.
 XX
 PR 30-OCT-2000; 2000US-244385P.
 XX
 PA (MAXY-) MAXYGEN INC.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
 PI Duck NB;
 XX
 DR N-PSDB; ABK92939.
 DR
 XX WPI; 2002-490010/52.
 PT
 PT New glyphosate-N-acetyltransferase gene and polypeptide having
 PT increased rate of catalysis and increased stability, useful for
 PT generating glyphosate resistant plants -
 PS
 PS Claim 127; Page 195; 254pp; English.
 XX
 CC The invention relates to an isolated or recombinant polypeptide (I)
 CC comprising an amino acid sequence that can be optimally aligned with
 CC a sequence (S) of 147 amino acids fully defined in the specification
 CC to generate a similarity score of at least 430, using the BLOSUM62
 CC matrix, a gap existence penalty of 11, and a gap extension penalty
 CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
 CC activity, including fragments, variants, polypeptides recognised by
 CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)
 CC (including variants and recombinant polynucleotides), a cell comprising
 CC the recombinant polynucleotide heterologous to the cell, a transgenic
 CC plant or its explant and seed comprising the above cell and producing (M)
 CC a variant of (I) by recursively recombining (I) with a second
 CC polynucleotide, forming a library of variant polynucleotides.
 CC The recombinant GAT polynucleotide is useful for producing a glyphosate
 CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
 CC These transgenic plants are useful for selectively controlling weeds in
 CC field containing a crop, and for producing crops such as cotton, corn or
 CC soybean. The transgenic plant is useful for controlling weeds in a field
 CC containing a crop and for preventing emergence of herbicide resistant
 CC weeds in a field. The present sequence represents a GAT polypeptide
 CC of the invention.
 CC

XX SQ Sequence 146 AA;
 Query Match 3.6%; Score 7; DB 23; Length 146;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 170 SSLIKHA 176
 |||||
 88 SSLIKHA 94
 DB

RESULT 43
 ABG63201
 ID ABG63201 standard; Protein; 146 AA.
 XX
 AC ABG63201;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Bacillus glyphosate N-acetyl transferase #246.
 XX
 KW Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
 KW N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
 KW transgenic; weed control; cotton; corn; soybean.
 OS
 XX Bacillus sp.
 XX
 PN MO200236782-A2.
 XX
 PD 10-MAY-2002.
 XX
 PF 29-OCT-2001; 2001MO-US46227.
 XX
 PR 30-OCT-2000; 2000US-244385P.
 XX
 PA (MAXY-) MAXYGEN INC.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Castle LA, Siehl D, Giver LJ, Minshull J, Ivy C, Chen YH;
 PI Duck NB;
 XX
 DR N-PSDB; ABK92947.
 DR
 XX WPI; 2002-490010/52.
 PT
 PT New glyphosate-N-acetyltransferase gene and polypeptide having
 PT increased rate of catalysis and increased stability, useful for
 PT generating glyphosate resistant plants -
 PS
 PS Claim 127; Page 196; 254pp; English.
 XX
 CC The invention relates to an isolated or recombinant polypeptide (I)
 CC comprising an amino acid sequence that can be optimally aligned with
 CC a sequence (S) of 147 amino acids fully defined in the specification
 CC to generate a similarity score of at least 430, using the BLOSUM62
 CC matrix, a gap existence penalty of 11, and a gap extension penalty
 CC of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
 CC activity, including fragments, variants, polypeptides recognised by
 CC anti-(S) antibodies. Also included are the polynucleotides encoding (I)
 CC (including variants and recombinant polynucleotides), a cell comprising
 CC the recombinant polynucleotide heterologous to the cell, a transgenic
 CC plant or its explant and seed comprising the above cell and producing (M)
 CC a variant of (I) by recursively recombining (I) with a second
 CC polynucleotide, forming a library of variant polynucleotides.
 CC The recombinant GAT polynucleotide is useful for producing a glyphosate
 CC (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
 CC These transgenic plants are useful for selectively controlling weeds in
 CC field containing a crop, and for producing crops such as cotton, corn or
 CC soybean. The transgenic plant is useful for controlling weeds in a field
 CC containing a crop and for preventing emergence of herbicide resistant
 CC weeds in a field. The present sequence represents a GAT polypeptide
 CC of the invention.
 CC

Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||||
88 SSLIKHA 94

RESULT 44
ABG63202
ID ABG63202 standard; Protein; 146 AA.

AC ABG63202;

DT 22-AUG-2002 (first entry)

DE Bacillus glyphosate N-acetyl transferase #247.

KM Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
KM transgenic; weed control; cotton; corn; soybean.

OS Bacillus sp.

PN WO200236782-A2.

PD 10-MAY-2002.

PF 29-OCT-2001; 2001WO-US46227.

PR 30-OCT-2000; 2000US-244385P.

PA (MAXY-) MAXYGEN INC.
PA (PION-) PIONEER HI-BRED INT INC.

PI Castle LA, Stiehl D, Giver LJ, Minshull J, Ivy C, Chen YH;

PI Duck NB;

DR WPI; 2002-490010/52.

DR N-PSDB; ABK92948.

PT New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants

PS Claim 127; Page 196; 254pp; English.

XX The invention relates to an isolated or recombinant polypeptide (I)
XX comprising an amino acid sequence that can be optimally aligned with
XX a sequence (S) of 147 amino acids fully defined in the specification
XX to generate a similarity score of at least 430, using the BLOSUM62
XX matrix, a gap existence penalty of 11, and a gap extension penalty
XX of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
XX activity, including fragments, variants, polypeptides recognised by
XX anti-(S) antibodies. Also included are the polynucleotides encoding (I)
XX (including variants and recombinant polynucleotides), a cell comprising
XX the recombinant polynucleotide heterologous to the cell, a transgenic
XX plant or its explant and seed comprising the above cell and producing (M)
XX a variant of (I) by recursively recombining (I) with a second
XX polynucleotide, forming a library of variant polynucleotides.
XX The recombinant GAT polynucleotide is useful for producing a glyphosate
XX (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
XX These transgenic plants are useful for selectively controlling weeds in
XX field containing a crop, and for producing crops such as cotton, corn or
XX soybean. The transgenic plant is useful for controlling weeds in a field
XX containing a crop and for preventing emergence of herbicide resistant
XX weeds in a field. The present sequence represents a GAT polypeptide
XX of the invention.

SQ Sequence 146 AA;

Query Match 3.6%; Score 7; DB 23; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||||
88 SSLIKHA 94

RESULT 45
ABG63204
ID ABG63204 standard; Protein; 146 AA.

AC ABG63204;

DT 22-AUG-2002 (first entry)

DE Bacillus glyphosate N-acetyl transferase #249.

KM Glyphosate resistance; glyphosate N-acetyltransferase; GAT;
KM N-phosphonomethylglycine; herbicide resistance; plant; enzyme;
KM transgenic; weed control; cotton; corn; soybean.

OS Bacillus sp.

PN WO200236782-A2.

PD 10-MAY-2002.

PF 29-OCT-2001; 2001WO-US46227.

PR 30-OCT-2000; 2000US-244385P.

PA (MAXY-) MAXYGEN INC.
PA (PION-) PIONEER HI-BRED INT INC.

PI Castle LA, Stiehl D, Giver LJ, Minshull J, Ivy C, Chen YH;

PI Duck NB;

DR WPI; 2002-490010/52.

DR N-PSDB; ABK92950.

PT New glyphosate-N-acetyltransferase gene and polypeptide having
PT increased rate of catalysis and increased stability, useful for
PT generating glyphosate resistant plants

PS Claim 127; Page 196; 254pp; English.

XX The invention relates to an isolated or recombinant polypeptide (I)
XX comprising an amino acid sequence that can be optimally aligned with
XX a sequence (S) of 147 amino acids fully defined in the specification
XX to generate a similarity score of at least 430, using the BLOSUM62
XX matrix, a gap existence penalty of 11, and a gap extension penalty
XX of 1, where the polypeptide has glyphosate N-acetyl transferase (GAT)
XX activity, including fragments, variants, polypeptides recognised by
XX anti-(S) antibodies. Also included are the polynucleotides encoding (I)
XX (including variants and recombinant polynucleotides), a cell comprising
XX the recombinant polynucleotide heterologous to the cell, a transgenic
XX plant or its explant and seed comprising the above cell and producing (M)
XX a variant of (I) by recursively recombining (I) with a second
XX polynucleotide, forming a library of variant polynucleotides.
XX The recombinant GAT polynucleotide is useful for producing a glyphosate
XX (N-phosphonomethylglycine) resistant transgenic plant or plant cell.
XX These transgenic plants are useful for selectively controlling weeds in
XX field containing a crop, and for producing crops such as cotton, corn or
XX soybean. The transgenic plant is useful for controlling weeds in a field
XX containing a crop and for preventing emergence of herbicide resistant
XX weeds in a field. The present sequence represents a GAT polypeptide
XX of the invention.

SQ Sequence 146 AA;

Query Match 3.6%; Score 7; DB 23; Length 146;

Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 170 SSLIKHA 176
Db 88 SSLIKHA 94

RESULT 46
AAU16056
ID AAU16056 standard; Protein: 151 AA.

AAU16056;

07-NOV-2001 (first entry)

Human novel secreted protein, Seq ID 1009.

Human; immunosuppressive; antiarthritic; antirheumatic;
cytotoxic; cardiac; vasotropic; cerebroprotective; nootropic;
neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
vulnerrary; secreted protein; rheumatoid arthritis;
hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
cerebrovascular disorder; cerebral ischaemia; angiogenesis;
nervous system disorder; Alzheimer's disease; infection; ocular disorder;
corneal infection; wound healing; epithelial cell proliferation;
skin ageing; food additive; preservative; antiproliferative.

Homo sapiens.

WO200155322-A2.

02-AUG-2001.

17-JAN-2001; 2001MO-US01341.

31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218280.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.

[illegible]

RESULT 47	
ABUS5125	ABUS5125 standard; Protein: 151 AA.
ID	ABUS5125
XX	ABUS5125;
AC	
XX	
DT	18-MAR-2003 (first entry)
XX	
DE	Human novel polypeptide #212.
XX	
KW	Human; neural disorder; immune system disorder; renal disorder; muscular disorder; respiratory disease; reproductive disorder; gastrointestinal disorder; pulmonary disorder; cardiovascular disorder; hyperproliferative disorder; inflammatory disease; allergic reaction
KW	blood related disorder; cancer; immunosuppressive; antiinflammatory; cardiovascular; nephrotropic; cytosatic; antiallergic; thrombolytic; haemostatic; antiarteriosclerotic.
KM	
XX	
OS	Homo sapiens.
XX	
PN	US2002132753-A1.
XX	
PD	19-SEP-2002.
XX	
PF	17-JAN-2001; 2001US-0764864.
XX	
PR	31-JAN-2000; 2000US-179065P.
PR	04-FEB-2000; 2000US-180628P.
PR	28-JUN-2000; 2000US-214886P.
PR	07-JUL-2000; 2000US-216647P.
PR	07-JUL-2000; 2000US-216880P.
PR	11-JUL-2000; 2000US-217487P.
PR	11-JUL-2000; 2000US-217466P.
PR	14-JUL-2000; 2000US-218290P.
PR	26-JUL-2000; 2000US-220963P.
PR	26-JUL-2000; 2000US-220966P.
PR	14-AUG-2000; 2000US-224518P.
PR	14-AUG-2000; 2000US-224519P.
PR	14-AUG-2000; 2000US-225267P.
PR	14-AUG-2000; 2000US-225268P.
PR	14-AUG-2000; 2000US-225270P.
PR	14-AUG-2000; 2000US-225447P.
PR	14-AUG-2000; 2000US-225757P.
PR	14-AUG-2000; 2000US-225758P.
PR	22-AUG-2000; 2000US-226866P.
PR	30-AUG-2000; 2000US-228924P.
PR	01-SEP-2000; 2000US-228927P.
PR	01-SEP-2000; 2000US-229343P.
PR	01-SEP-2000; 2000US-229344P.
PR	01-SEP-2000; 2000US-229345P.
PR	05-SEP-2000; 2000US-229503P.
PR	05-SEP-2000; 2000US-229513P.
PR	08-SEP-2000; 2000US-229513P.
PR	21-SEP-2000; 2000US-234223P.
PR	21-SEP-2000; 2000US-234274P.
PR	25-SEP-2000; 2000US-234997P.
PR	27-SEP-2000; 2000US-235834P.
PR	29-SEP-2000; 2000US-236327P.
PR	29-SEP-2000; 2000US-236367P.
PR	29-SEP-2000; 2000US-236368P.
PR	29-SEP-2000; 2000US-236370P.
PR	02-OCT-2000; 2000US-236802P.
PR	02-OCT-2000; 2000US-237037P.
PR	02-OCT-2000; 2000US-237038P.
PR	02-OCT-2000; 2000US-237039P.
PR	13-OCT-2000; 2000US-237040P.
PR	13-OCT-2000; 2000US-239335P.
PR	20-OCT-2000; 2000US-240960P.
PR	20-OCT-2000; 2000US-241785P.
PR	20-OCT-2000; 2000US-241809P.
PR	01-NOV-2000; 2000US-244617P.

PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251856P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI, 2003-147444/14.
XX N-PSDB; ABX73384.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for
PT treating, inhibiting or preventing e.g. neural, immune system,
PT muscular, respiratory, reproductive, gastrointestinal, pulmonary,
PT cardiovascular or renal disorders -
XX
XX Claim 11, SEQ ID NO 1009; 402bp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid
CC arthritis and multiple sclerosis), muscular disorders, respiratory
CC diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis),
CC reproductive disorders, gastrointestinal disorders, pulmonary disorders,
CC cardiovascular disorders (e.g. congenital heart defects, Epstein's
CC anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute
CC kidney failure and end-stage renal disease), hyperproliferative disorders
CC (e.g. Hodgkin's disease and leukemia), inflammatory diseases (e.g.
CC septic shock, burstitis and appendicitis), allergic reactions and
CC conditions (e.g. asthma), blood related disorders (e.g. thrombosis,
CC atherosclerosis and myocardial infarction) and cancerous diseases.
CC Sequences ABUS4914-ABUS5699 and ABUS5748 represent human novel
CC polypeptides of the invention.
XX
XX Sequence 151 AA;
SO

Query Match 3.6%; Score 7; DB 24; Length 151;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SSSLIRK 175
Db 134 SSSLIRK 140

RESULT 48
AAG18890
ID AAG18890 standard; Protein; 152 AA.
XX
AC AAG18890;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 20479.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135829.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147312.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148568.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149375.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 3.6%; Score 7; DB 21; Length 152;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 FFFLLPI 10
 Db 4 FFFLLPI 10

RESULT 49
 ABG72337
 ID ABG72337 standard; Protein; 160 AA.
 AC ABG72337;
 XX 31-JAN-2003 (first entry)
 DT 31-JAN-2003 (first entry)
 XX Human hypoxin inductive factor 1 alpha subunit 17.60.
 DE Human hypoxin inductive factor 1 alpha subunit 17.60;
 XX Human; hypoxin inductive factor 1 alpha subunit 17.60;
 KW erythrocyte disease; haemopathy; substance metabolic disorder; tumour;
 KM human immunodeficiency virus infection; HIV; immunological disease;
 KW inflammation.
 XX Homo sapiens.
 OS Homo sapiens.
 XX CN1352128-A.
 PN CN1352128-A.
 XX 05-JUN-2002.
 PD 05-JUN-2002.
 XX 06-NOV-2000; 2000CN-0127266.
 PF 06-NOV-2000; 2000CN-0127266.
 XX 06-NOV-2000; 2000CN-0127266.
 PR 06-NOV-2000; 2000CN-0127266.
 XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 XX Mao Y, Xie Y;
 PI Mao Y, Xie Y;
 XX WPI; 2002-637136/69.
 DR N-PSDB; ABS57125.

XX New human hypoxic inductive factor 1 alpha subunit 17.60 polypeptide
 PT for treating e.g. erythrocyte disease, hemopathy, various tumors, human
 PT immunodeficiency virus infection, immunological diseases, and
 PT inflammation -
 XX
 XX Claim 1; Page 28 (disclosure): 34pp; Chinese.
 XX
 CC The present invention discloses a new kind of polypeptide, human hypoxin
 CC inductive factor 1 alpha subunit 17.60, polynucleotides encoding the
 CC polypeptide and production of the protein using recombinant DNA
 CC technology. The present invention also discloses applying the
 CC polypeptide in treating various diseases, such as erythrocyte disease,
 CC hemopathy, substance metabolic disorder, various tumors, human
 CC immunodeficiency virus (HIV) infection, immunological diseases and
 CC inflammations. The present invention also discloses the antagonist
 CC resisting the polypeptide and its treatment effect. The present
 CC invention also discloses application of polynucleotides encoding human
 CC hypoxic inductive factor 1 alpha subunit 17.60. The present
 CC sequence represents human hypoxic inductive factor 1 alpha subunit 17.60.
 CC
 SO Sequence 160 AA;
 Query Match 3.6%; Score 7; DB 23; Length 160;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 146 ISLESSG 152
 DB 27 ISLESSG 33
 RESULT 50
 ABB11250
 ID ABB11250 standard; peptide; 162 AA.
 XX
 XX ABB11250;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 XX Human Fe-gamma receptor homologue, SEQ ID NO:1620.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemokinesis; chemokinesis; thrombolytic; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytoactive; osteopathic; vasotropic; cardiac; virucide; antibacterial;
 KW antifungal; vulnery; antitumor.
 XX
 OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PP 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT,
 XX
 DR WPI, 2001-457740/49.
 DR N-PSDB; ABA08494.

XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 XX Claim 20; Page 160; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 CC
 SO Sequence 162 AA;
 Query Match 3.6%; Score 7; DB 22; Length 162;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 DRKHALE 69
 DB 94 DRKHALE 100

Search completed: January 28, 2004, 18:20:18
 Job time : 75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: January 28, 2004, 18:20:25 ; Search time 602 Seconds
(without alignments)
66.279 Million cell updates/sec

Title: US-09-881-556a-4

Perfect score: 193

Sequence: 1 HEHFFLPRLPISIFPDSNDA.....KHAPIKNGHNNQSEKSKYK 193

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 777136 seqs, 206736638 residues

Word size : 5

Total number of hits satisfying chosen parameters: 25128

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	4.7	19662	12 US-10-084-846a-6	Sequence 6, Appl1
2	7	3.6	72	12 US-10-319-763-93	Sequence 93, Appl
3	7	3.6	72	12 US-10-319-763-187	Sequence 187, Appl
4	7	3.6	100	9 US-09-853-161-61	Sequence 61, Appl
5	7	3.6	100	9 US-09-852-659a-61	Sequence 61, Appl
6	7	3.6	100	10 US-09-852-797-61	Sequence 61, Appl
7	7	3.6	106	12 US-10-319-763-95	Sequence 95, Appl
8	7	3.6	106	12 US-10-319-763-189	Sequence 189, Appl
9	7	3.6	146	15 US-10-004-357-440	Sequence 440, App
10	7	3.6	146	15 US-10-004-357-441	Sequence 441, App
11	7	3.6	146	15 US-10-004-357-445	Sequence 445, App
12	7	3.6	146	15 US-10-004-357-452	Sequence 452, App
13	7	3.6	146	15 US-10-004-357-457	Sequence 457, App
14	7	3.6	146	15 US-10-004-357-458	Sequence 458, App
15	7	3.6	146	15 US-10-004-357-462	Sequence 462, App

16	7	3.6	146	15 US-10-004-357-464	Sequence 464, App
17	7	3.6	146	15 US-10-004-357-465	Sequence 465, App
18	7	3.6	146	15 US-10-004-357-470	Sequence 470, App
19	7	3.6	146	15 US-10-004-357-473	Sequence 473, App
20	7	3.6	146	15 US-10-004-357-485	Sequence 485, App
21	7	3.6	146	15 US-10-004-357-487	Sequence 487, App
22	7	3.6	146	15 US-10-004-357-488	Sequence 488, App
23	7	3.6	146	15 US-10-004-357-489	Sequence 489, App
24	7	3.6	146	15 US-10-004-357-490	Sequence 490, App
25	7	3.6	146	15 US-10-004-357-491	Sequence 491, App
26	7	3.6	146	15 US-10-004-357-492	Sequence 492, App
27	7	3.6	146	15 US-10-004-357-495	Sequence 495, App
28	7	3.6	146	15 US-10-004-357-503	Sequence 503, App
29	7	3.6	146	15 US-10-004-357-504	Sequence 504, App
30	7	3.6	146	15 US-10-004-357-506	Sequence 506, App
31	7	3.6	151	10 US-09-764-864-1809	Sequence 1809, App
32	7	3.6	212	10 US-09-902-941-1820	Sequence 1820, App
33	7	3.6	212	10 US-09-902-941-1820	Sequence 1820, App
34	7	3.6	212	10 US-09-849-626-1820	Sequence 1820, App
35	7	3.6	212	10 US-10-113-872-1820	Sequence 1820, App
36	7	3.6	212	10 US-10-017-754-1820	Sequence 1820, App
37	7	3.6	294	9 US-09-838-626-6320	Sequence 6320, App
38	7	3.6	362	12 US-10-369-493-442	Sequence 442, App
39	7	3.6	374	12 US-10-108-260A-3171	Sequence 3171, App
40	7	3.6	397	12 US-10-183-708-8	Sequence 8, Appl1
41	7	3.6	397	12 US-09-933-227-8	Sequence 8, Appl1
42	7	3.6	429	10 US-09-738-626-5318	Sequence 5318, App
43	7	3.6	429	10 US-09-934-455-170	Sequence 170, App
44	7	3.6	489	11 US-10-225-068-166	Sequence 166, App
45	7	3.6	495	12 US-10-029-366-34057	Sequence 34057, A
46	7	3.6	537	12 US-10-369-493-15951	Sequence 15951, A
47	7	3.6	537	12 US-10-369-493-16315	Sequence 16315, A
48	7	3.6	627	14 US-10-047-260-36	Sequence 2798, App
49	7	3.6	627	14 US-10-047-260-36	Sequence 36, Appl
50	7	3.6	784	12 US-10-094-749-2464	Sequence 2464, App
51	7	3.6	909	12 US-10-369-493-16893	Sequence 16893, A
52	7	3.6	1693	8 US-10-114-170-257	Sequence 257, App
53	7	3.1	8	8 US-08-817-8328-26	Sequence 26, Appl
54	6	3.1	14	9 US-09-823-936-46	Sequence 46, Appl
55	6	3.1	14	14 US-10-193-142-21	Sequence 21, Appl
56	6	3.1	14	15 US-10-134-493-21	Sequence 21, Appl
57	6	3.1	14	15 US-10-061-216-21	Sequence 21, Appl
58	6	3.1	14	15 US-10-061-216-21	Sequence 21, Appl
59	6	3.1	14	15 US-10-286-140-46	Sequence 46, Appl
60	6	3.1	15	9 US-09-823-936-45	Sequence 45, Appl
61	6	3.1	15	14 US-10-193-142-20	Sequence 20, Appl
62	6	3.1	15	15 US-10-134-493-20	Sequence 20, Appl
63	6	3.1	15	15 US-10-061-216-20	Sequence 20, Appl
64	6	3.1	15	15 US-10-286-140-45	Sequence 45, Appl
65	6	3.1	18	9 US-09-823-936-44	Sequence 44, Appl
66	6	3.1	18	14 US-10-193-142-19	Sequence 19, Appl
67	6	3.1	18	15 US-10-134-493-19	Sequence 19, Appl
68	6	3.1	18	15 US-10-061-216-19	Sequence 19, Appl
69	6	3.1	18	15 US-10-286-140-44	Sequence 44, Appl
70	6	3.1	19	9 US-09-823-936-43	Sequence 43, Appl
71	6	3.1	19	14 US-10-193-142-18	Sequence 18, Appl
72	6	3.1	19	15 US-10-134-493-18	Sequence 18, Appl
73	6	3.1	19	15 US-10-061-216-18	Sequence 18, Appl
74	6	3.1	19	15 US-10-286-140-43	Sequence 43, Appl
75	6	3.1	20	9 US-09-823-936-42	Sequence 42, Appl
76	6	3.1	20	14 US-10-193-142-17	Sequence 17, Appl
77	6	3.1	20	15 US-10-134-493-17	Sequence 17, Appl
78	6	3.1	20	15 US-10-061-216-17	Sequence 17, Appl
79	6	3.1	20	15 US-10-286-140-42	Sequence 42, Appl
80	6	3.1	21	9 US-09-823-936-41	Sequence 41, Appl
81	6	3.1	21	14 US-10-193-142-16	Sequence 16, Appl
82	6	3.1	21	15 US-10-134-493-16	Sequence 16, Appl
83	6	3.1	21	15 US-10-061-216-16	Sequence 16, Appl
84	6	3.1	21	15 US-10-286-140-41	Sequence 41, Appl
85	6	3.1	23	9 US-09-823-936-40	Sequence 40, Appl
86	6	3.1	23	14 US-10-193-142-15	Sequence 15, Appl
87	6	3.1	23	15 US-10-134-493-15	Sequence 15, Appl
88	6	3.1	23	15 US-10-061-216-15	Sequence 15, Appl

89	6	3.1	23	15	US-10-286-140-40	Sequence 40, Appl
90	6	3.1	24	9	US-09-823-936-39	Sequence 39, Appl
91	6	3.1	24	15	US-10-193-142-14	Sequence 14, Appl
92	6	3.1	24	15	US-10-134-493-14	Sequence 14, Appl
93	6	3.1	24	15	US-10-061-216-14	Sequence 14, Appl
94	6	3.1	24	15	US-10-286-140-39	Sequence 39, Appl
95	6	3.1	26	9	US-09-864-761-38522	Sequence 38522, A
96	6	3.1	37	10	US-09-972-484-20	Sequence 20, Appl
97	6	3.1	39	12	US-10-029-386-29390	Sequence 29390, A
98	6	3.1	42	9	US-09-864-761-41263	Sequence 41263, A
99	6	3.1	46	15	US-10-201-444-5	Sequence 5, Appl
100	6	3.1	48	9	US-09-864-761-44568	Sequence 44568, A

ALIGNMENTS

RESULT 1
US-10-084-846a-6
; Sequence 6, Application US/10084846a
; Publication No. US2004006026A1
; GENERAL INFORMATION:
; APPLICANT: WEINBAUER, GABRIELE
; APPLICANT: MOHLEWEG, AGNES
; APPLICANT: TREFFER, AXEL
; APPLICANT: BECHTOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084, 846a
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patentin Ver. 3.2
; SEQ ID NO 6
; LENGTH: 19662
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 2.
; OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846a-6

Query Match 4.7%; Score 9; DB 12; Length 19662;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 150 SSGENLIST 158
DB 19533 SSGENLIST 19541

RESULT 2
US-10-319-763-93
; Sequence 93, Application US/10319763
; Publication No. US20030144490A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Ducleert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: G-031.US04.DIV
; CURRENT APPLICATION NUMBER: US/10/319, 763
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/066, 677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069, 957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074, 121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081, 563

PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096, 116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099, 273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 93
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -32...-1
US-10-319-763-93

Query Match 3.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 NKISLES 150
DB 59 NKISLES 65

RESULT 3
US-10-319-763-187
; Sequence 187, Application US/10319763
; Publication No. US20030144490A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Ducleert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: G-031.US04.DIV
; CURRENT APPLICATION NUMBER: US/10/319, 763
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/066, 677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069, 957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074, 121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081, 563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096, 116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099, 273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 187
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -32...-1
US-10-319-763-187

Query Match 3.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 NKISLES 150
DB 59 NKISLES 65

RESULT 4
US-09-853-161-61
; Sequence 61, Application US/09853161
; Patent No. US20020076756A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 28 Human Secreted Proteins
/ FILE REFERENCE: P2003P3
/ CURRENT APPLICATION NUMBER: US/09/853,161
/ CURRENT FILING DATE: 2001-05-11
/ PRIOR APPLICATION NUMBER: 60/265,583
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 09/152,060
/ PRIOR FILING DATE: 1998-09-11
/ PRIOR APPLICATION NUMBER: PCT/US98/04858
/ PRIOR FILING DATE: 1998-03-12
/ PRIOR APPLICATION NUMBER: 60/040,762
/ PRIOR FILING DATE: 1997-03-14
/ PRIOR APPLICATION NUMBER: 60/040,710
/ PRIOR FILING DATE: 1997-03-14
/ PRIOR APPLICATION NUMBER: 60/050,934
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: 60/048,100
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: 60/048,357
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: 60/048,189
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: 60/057,765
/ PRIOR FILING DATE: 1997-09-05
/ PRIOR APPLICATION NUMBER: 60/048,970
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 60/068,368
/ PRIOR FILING DATE: 1997-12-19
/ NUMBER OF SEQ ID NOS: 118
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO: 61
/ LENGTH: 100
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-853-161-61

Query Match      3.6%  Score 7;  DB 9;  Length 100;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
/ PRIOR APPLICATION NUMBER: 60/048,189
/ PRIOR FILING DATE: 1997-05-30
/ PRIOR APPLICATION NUMBER: 60/057,765
/ PRIOR FILING DATE: 1997-09-05
/ PRIOR APPLICATION NUMBER: 60/048,970
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 60/068,368
/ PRIOR FILING DATE: 1997-12-19
/ NUMBER OF SEQ ID NOS: 121
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO: 61
/ LENGTH: 100
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-852-659A-61

Query Match      3.6%  Score 7;  DB 9;  Length 100;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 53 NKISLES 59

RESULT 7
US-10-319-763-95; Sequence 95, Application US/10319763
; Publication No. US20030144490A1

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean-Baptiste

; APPLICANT: Duclert, Aymeric

; APPLICANT: Bougueleret, Lydie

; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS

; FILE REFERENCE: G-031.US04.DIV

; CURRENT APPLICATION NUMBER: US/10/319,763

; CURRENT FILING DATE: 2002-12-10

; PRIOR APPLICATION NUMBER: 60/066,677

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/069,957

; PRIOR FILING DATE: 1997-12-17

; PRIOR APPLICATION NUMBER: 60/074,121

; PRIOR FILING DATE: 1998-02-09

; PRIOR APPLICATION NUMBER: 60/081,563

; PRIOR FILING DATE: 1998-04-13

; PRIOR APPLICATION NUMBER: 60/096,116

; PRIOR FILING DATE: 1998-08-10

; PRIOR APPLICATION NUMBER: 60/099,273

; PRIOR FILING DATE: 1998-09-04

; NUMBER OF SEQ ID NOS: 229

; SOFTWARE: Patent.pm

; SEQ ID NO 95

; LENGTH: 106

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: -32...-1

US-10-319-763-95

Query Match 3.6%; Score 7; DB 12; Length 106;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 144 NKISLES 150

Db 59 NKISLES 65

RESULT 8
US-10-319-763-189

; Sequence 189, Application US/10319763

; Publication No. US20030144490A1

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean-Baptiste

; APPLICANT: Duclert, Aymeric

; APPLICANT: Bougueleret, Lydie

; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS

; FILE REFERENCE: G-031.US04.DIV

; CURRENT APPLICATION NUMBER: US/10/319,763

; CURRENT FILING DATE: 2002-12-10

; PRIOR APPLICATION NUMBER: 60/066,677

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/069,957

; PRIOR FILING DATE: 1997-12-17

; PRIOR APPLICATION NUMBER: 60/074,121

; PRIOR FILING DATE: 1998-02-09

; PRIOR APPLICATION NUMBER: 60/081,563

; PRIOR FILING DATE: 1998-04-13

; PRIOR APPLICATION NUMBER: 60/096,116

; PRIOR FILING DATE: 1998-08-10

; PRIOR APPLICATION NUMBER: 60/099,273

; NUMBER OF SEQ ID NOS: 229

; SOFTWARE: Patent.pm

; SEQ ID NO 189

; LENGTH: 106

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: -32...-1

US-10-319-763-189

Query Match 3.6%; Score 7; DB 12; Length 106;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 144 NKISLES 150

Db 59 NKISLES 65

RESULT 9
US-10-004-357-440

; Sequence 440, Application US/10004357

; Publication No. US20030083480A1

; GENERAL INFORMATION:

; APPLICANT: Caetle, Linda A.

; APPLICANT: Siehl, Dan

; APPLICANT: Giver, Lorraine

; APPLICANT: Minshull, Jeremy

; APPLICANT: Ivy, Christina

; APPLICANT: Chen, Yong Hong

; APPLICANT: Duck, Nicholas B.

; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE

; FILE REFERENCE: 02-107010US

; CURRENT APPLICATION NUMBER: US/10/004,357

; CURRENT FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: US 60/244,385

; PRIOR FILING DATE: 2000-10-30

; NUMBER OF SEQ ID NOS: 515

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 440

; LENGTH: 146

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic Protein Sequence

US-10-004-357-440

Query Match 3.6%; Score 7; DB 15; Length 146;

Best Local Similarity 100.0%; Pred. No. 82;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 170 SSLIKHA 176

Db 88 SSLIKHA 94

RESULT 10
US-10-004-357-441

; Sequence 441, Application US/10004357

; Publication No. US20030083480A1

; GENERAL INFORMATION:

; APPLICANT: Caetle, Linda A.

; APPLICANT: Siehl, Dan

; APPLICANT: Giver, Lorraine

; APPLICANT: Minshull, Jeremy

; APPLICANT: Ivy, Christina

; APPLICANT: Chen, Yong Hong

; APPLICANT: Duck, Nicholas B.

; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE

; FILE REFERENCE: 02-107010US

; CURRENT APPLICATION NUMBER: US/10/004,357

; CURRENT FILING DATE: 2001-10-29

PRIOR APPLICATION NUMBER: US 60/244,385
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 515
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 441
LENGTH: 146
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-441

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
DB 88 SSLIKHA 94

RESULT 11
US-10-004-357-445
Sequence 445, Application US/10004357
Publication No. US20030083480A1
GENERAL INFORMATION:
APPLICANT: Castile, Linda A.
APPLICANT: Givert, Lorraine
APPLICANT: Minshull, Jeremy
APPLICANT: Ivy, Christina
APPLICANT: Chen, Yong Hong
APPLICANT: Duck, Nicholas B.
TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
TITLE OF INVENTION: (GAT) GENES
FILE REFERENCE: 02-107010US
CURRENT APPLICATION NUMBER: US/10/004,357
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: US 60/244,385
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 515
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 445
LENGTH: 146
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-445

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
DB 88 SSLIKHA 94

RESULT 12
US-10-004-357-452
Sequence 452, Application US/10004357
Publication No. US20030083480A1
GENERAL INFORMATION:
APPLICANT: Castile, Linda A.
APPLICANT: Siehl, Dan
APPLICANT: Givert, Lorraine
APPLICANT: Minshull, Jeremy
APPLICANT: Ivy, Christina
APPLICANT: Chen, Yong Hong
APPLICANT: Duck, Nicholas B.
TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
TITLE OF INVENTION: (GAT) GENES

FILE REFERENCE: 02-107010US
CURRENT APPLICATION NUMBER: US/10/004,357
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: US 60/244,385
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 515
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 452
LENGTH: 146
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Protein Sequence
NAME/KEY: unsure
LOCATION: 45
OTHER INFORMATION: Xaa = Phe, Ser, Tyr, or Cys
NAME/KEY: unsure
LOCATION: 105
OTHER INFORMATION: Xaa = Leu, Met, or Val
US-10-004-357-452

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
DB 88 SSLIKHA 94

RESULT 13
US-10-004-357-457
Sequence 457, Application US/10004357
Publication No. US20030083480A1
GENERAL INFORMATION:
APPLICANT: Castile, Linda A.
APPLICANT: Siehl, Dan
APPLICANT: Givert, Lorraine
APPLICANT: Minshull, Jeremy
APPLICANT: Ivy, Christina
APPLICANT: Chen, Yong Hong
APPLICANT: Duck, Nicholas B.
TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
TITLE OF INVENTION: (GAT) GENES
FILE REFERENCE: 02-107010US
CURRENT APPLICATION NUMBER: US/10/004,357
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: US 60/244,385
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 515
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 457
LENGTH: 146
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-457

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
DB 88 SSLIKHA 94

RESULT 14
US-10-004-357-458
Sequence 458, Application US/10004357
Publication No. US20030083480A1
GENERAL INFORMATION:

APPLICANT: Caetle, Linda A.
APPLICANT: Siehl, Dan
APPLICANT: Giver, Lorraine
APPLICANT: Minshull, Jeremy
APPLICANT: Ivy, Christina
APPLICANT: Chen, Yong Hong
APPLICANT: Duck, Nicholas B.
TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
FILE REFERENCE: 02-107010US
CURRENT APPLICATION NUMBER: US/10/004,357
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: US 60/244,385
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 515
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 458
LENGTH: 146
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-458

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 SSLIKHA 176
|||||
Db 88 SSLIKHA 94

RESULT 15
US-10-004-357-462
Sequence 462, Application US/10004357
Publication No. US20030083480A1
GENERAL INFORMATION:
APPLICANT: Caetle, Linda A.
APPLICANT: Siehl, Dan
APPLICANT: Giver, Lorraine
APPLICANT: Minshull, Jeremy
APPLICANT: Ivy, Christina
APPLICANT: Chen, Yong Hong
APPLICANT: Duck, Nicholas B.
TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
FILE REFERENCE: 02-107010US
CURRENT APPLICATION NUMBER: US/10/004,357
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: US 60/244,385
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 515
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 462
LENGTH: 146
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-462

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 SSLIKHA 176
|||||
Db 88 SSLIKHA 94

RESULT 16
US-10-004-357-464

Sequence 464, Application US/10004357
Publication No. US20030083480A1
GENERAL INFORMATION:
APPLICANT: Caetle, Linda A.
APPLICANT: Siehl, Dan
APPLICANT: Giver, Lorraine
APPLICANT: Minshull, Jeremy
APPLICANT: Ivy, Christina
APPLICANT: Chen, Yong Hong
APPLICANT: Duck, Nicholas B.
TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
FILE REFERENCE: 02-107010US
CURRENT APPLICATION NUMBER: US/10/004,357
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: US 60/244,385
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 515
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 464
LENGTH: 146
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-464

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 SSLIKHA 176
|||||
Db 88 SSLIKHA 94

RESULT 17
US-10-004-357-465
Sequence 465, Application US/10004357
Publication No. US20030083480A1
GENERAL INFORMATION:
APPLICANT: Caetle, Linda A.
APPLICANT: Siehl, Dan
APPLICANT: Giver, Lorraine
APPLICANT: Minshull, Jeremy
APPLICANT: Ivy, Christina
APPLICANT: Chen, Yong Hong
APPLICANT: Duck, Nicholas B.
TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
FILE REFERENCE: 02-107010US
CURRENT APPLICATION NUMBER: US/10/004,357
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: US 60/244,385
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 515
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 465
LENGTH: 146
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-465

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 SSLIKHA 176
|||||
Db 88 SSLIKHA 94

RESULT 18
US-10-004-357-470
; Sequence 470, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 470
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-470

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||||
DB 88 SSLIKHA 94

RESULT 19
US-10-004-357-473
; Sequence 473, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Giver, Lorraine
; APPLICANT: Siehl, Dan
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 473
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-473

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLIKHA 176

DB 88 SSLIKHA 94
|||||
RESULT 20
US-10-004-357-485
; Sequence 485, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 485
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-485

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||||
DB 88 SSLIKHA 94

RESULT 21
US-10-004-357-487
; Sequence 487, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Giver, Lorraine
; APPLICANT: Siehl, Dan
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 487
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-487

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SSLIKHA 176
|||||||
Db 88 SSLIKHA 94

RESULT 22
US-10-004-357-488
; Sequence 488, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-488

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||||||
Db 88 SSLIKHA 94

RESULT 23
US-10-004-357-489
; Sequence 489, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 489
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-489

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||||||
Db 88 SSLIKHA 94

RESULT 24
US-10-004-357-490
; Sequence 490, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 490
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-490

Query Match 3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKHA 176
|||||||
Db 88 SSLIKHA 94

RESULT 25
US-10-004-357-491
; Sequence 491, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 491
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-491
Query Match      3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      170 SSLIKHA 176
      |||||
Db      88 SSLIKHA 94

RESULT 26
US-10-004-357-492
; Sequence 492, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Chen, Yong Hong
; APPLICANT: Ivy, Christina
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 492
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-492
Query Match      3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      170 SSLIKHA 176
      |||||
Db      88 SSLIKHA 94

RESULT 27
US-10-004-357-495
; Sequence 495, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 495
```

```

; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
; NAME/KEY: unsure
; LOCATION: 3
; OTHER INFORMATION: Xaa = Asp or Glu
; NAME/KEY: unsure
; LOCATION: 26
; OTHER INFORMATION: Xaa = Leu, Ile, or Val
; NAME/KEY: 33
; LOCATION: 33
; OTHER INFORMATION: Xaa = Ile, Thr, Asn, or Ser
US-10-004-357-495
Query Match      3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      170 SSLIKHA 176
      |||||
Db      88 SSLIKHA 94

RESULT 28
US-10-004-357-503
; Sequence 503, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 503
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
US-10-004-357-503
Query Match      3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      170 SSLIKHA 176
      |||||
Db      88 SSLIKHA 94

RESULT 29
US-10-004-357-504
; Sequence 504, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Castle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
```

```

; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTURANSFERASE
; TITLE OF INVENTION: (GAT) GENES
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 504
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
; US-10-004-357-504

```

```

Query Match          3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      170 SSLIKHA 176
      |||||
Db      88 SSLIKHA 94

```

```

RESULT 30
; US-10-004-357-506
; Sequence 506, Application US/10004357
; Publication No. US20030083480A1
; GENERAL INFORMATION:
; APPLICANT: Cattle, Linda A.
; APPLICANT: Siehl, Dan
; APPLICANT: Giver, Lorraine
; APPLICANT: Minshull, Jeremy
; APPLICANT: Ivy, Christina
; APPLICANT: Chen, Yong Hong
; APPLICANT: Duck, Nicholas B.
; TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTURANSFERASE
; TITLE OF INVENTION: (GAT) GENES
; FILE REFERENCE: 02-107010US
; CURRENT APPLICATION NUMBER: US/10/004,357
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/244,385
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 506
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Protein Sequence
; US-10-004-357-506

```

```

Query Match          3.6%; Score 7; DB 15; Length 146;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      170 SSLIKHA 176
      |||||
Db      88 SSLIKHA 94

```

```

RESULT 31
; US-09-764-864-1009
; Sequence 1009, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```

```

; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1009
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (88)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-864-1009

```

```

Query Match          3.6%; Score 7; DB 10; Length 151;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      169 SSSLIKH 175
      |||||
Db      134 SSSLIKH 140

```

```

RESULT 32
; US-09-736-457-1820
; Sequence 1820, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Lodges, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1820
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-736-457-1820

```

```

Query Match          3.6%; Score 7; DB 10; Length 212;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      84 CFEKNEA 90
      |||||
Db      121 CFEKNEA 127

```

```

RESULT 33
; US-09-902-941-1820
; Sequence 1820, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnetakis, Margarita

```

```

/ APPLICANT: Carter, Darrick
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: McNabb, Andria
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 210121.478C17
/ CURRENT APPLICATION NUMBER: US/09/902,941
/ CURRENT FILING DATE: 2001-07-10
/ NUMBER OF SEQ ID NOS: 2002
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1820
/ LENGTH: 212
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-902-941-1820

```

```

Query Match      3.6%; Score 7; DB 10; Length 212;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      84 CFEKNEA 90
        |||||
Db      121 CFEKNEA 127

```

```

RESULT 34
US-09-849-626-1820
/ Sequence 1820, Application US/09849626
/ Publication No. US20020197669A1
/ GENERAL INFORMATION:
/ APPLICANT: Bangur, Chaitanya
/ APPLICANT: Fanger, Gary
/ APPLICANT: Wang, Aijun
/ APPLICANT: Wang, Tonglong
/ APPLICANT: Switzer, Anne
/ APPLICANT: McNeill, Patricia
/ APPLICANT: Clapper, Jonathan
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C16
/ CURRENT APPLICATION NUMBER: US/09/849,626
/ CURRENT FILING DATE: 2001-05-03
/ NUMBER OF SEQ ID NOS: 1926
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1820
/ LENGTH: 212
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-849-626-1820

```

```

Query Match      3.6%; Score 7; DB 10; Length 212;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      84 CFEKNEA 90
        |||||
Db      121 CFEKNEA 127

```

```

RESULT 35
US-10-113-872-1820
/ Sequence 1820, Application US/10113872
/ Publication No. US20030170255A1
/ GENERAL INFORMATION:
/ APPLICANT: Watanabe, Yoshihiro
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Sleach, Paul R.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Fanger, Gary R.

```

```

/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 210121.478C19
/ CURRENT APPLICATION NUMBER: US/10/113,872
/ CURRENT FILING DATE: 2002-03-28
/ NUMBER OF SEQ ID NOS: 2011
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1820
/ LENGTH: 212
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-113-872-1820

```

```

Query Match      3.6%; Score 7; DB 12; Length 212;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      84 CFEKNEA 90
        |||||
Db      121 CFEKNEA 127

```

```

RESULT 36
US-10-017-754-1820
/ Sequence 1820, Application US/10017754
/ Publication No. US20030054363A1
/ GENERAL INFORMATION:
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Wang, Tonglong
/ APPLICANT: Watanabe, Yoshihiro
/ APPLICANT: Johnson, Jeffrey C.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Marnerakis, Margarita
/ APPLICANT: Carter, Darrick
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: McNabb, Andria
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 210121.478C18
/ CURRENT APPLICATION NUMBER: US/10/017,754
/ CURRENT FILING DATE: 2001-10-29
/ NUMBER OF SEQ ID NOS: 2004
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1820
/ LENGTH: 212
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-017-754-1820

```

```

Query Match      3.6%; Score 7; DB 15; Length 212;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      84 CFEKNEA 90
        |||||
Db      121 CFEKNEA 127

```

```

RESULT 37
US-09-838-564A-3
/ Sequence 3, Application US/09838564A
/ Patent No. US20020072099A1
/ GENERAL INFORMATION:
/ APPLICANT: Mockel, Bettina
/ APPLICANT: Pfeifferle, Walter
/ APPLICANT: Marx, Achim
/ TITLE OF INVENTION: New Nucleotide Sequences Coding for the Genes succ and succD
/ FILE REFERENCE: 21123/278466
/ CURRENT APPLICATION NUMBER: US/09/838,564A
/ CURRENT FILING DATE: 2001-04-20
/ NUMBER OF SEQ ID NOS: 5

```

```
SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-838-564A-3
```

Query Match 3.6%; Score 7; DB 9; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 KHALEAV 71
|||||
Db 266 KHALEAV 272

```
RESULT 38
US-09-738-626-6320
; Sequence 6320, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OGISHI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6320
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6320
```

Query Match 3.6%; Score 7; DB 10; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 KHALEAV 71
|||||
Db 266 KHALEAV 272

```
RESULT 39
US-10-369-493-442
; Sequence 442, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
```

```
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 442
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-442
```

Query Match 3.6%; Score 7; DB 12; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 KSLFVLN 27
|||||
Db 76 KSLFVLN 82

```
RESULT 40
US-10-108-260A-3171
; Sequence 3171, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 3171
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3171
```

Query Match 3.6%; Score 7; DB 12; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 DRKHALE 69
|||||
Db 306 DRKHALE 312

```
RESULT 41
US-10-183-708-8
; Sequence 8, Application US/10183708
; Publication No. US20030143679A1
; GENERAL INFORMATION:
; APPLICANT: VOSSHALL, LESLIE
; APPLICANT: AMREIN, HUBERT
; APPLICANT: AXEL, RICHARD
; TITLE OF INVENTION: GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF
; FILE REFERENCE: 0575/58715-AA-PCT-US/JPW/ADM/BJA
; CURRENT APPLICATION NUMBER: US/10/183,708
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 09/932,227
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/04995
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/257,706
; PRIOR FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 397
; TYPE: PRT
; ORGANISM: DROSOPHILA MELANOGASTER
US-10-183-708-8
```

Query Match 3.6%; Score 7; DB 12; Length 397;
Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 LIPISIS 13
|||||
Db 62 LIPISIS 68

RESULT 42
US-09-932-227-8
; Sequence 8, Application US/09932227
; Publication No. US20030186359A1
; GENERAL INFORMATION:
; APPLICANT: VOSSSHALL, LESLIE
; APPLICANT: AMREIN, HUBERT
; APPLICANT: AXEL, RICHARD
; TITLE OF INVENTION: GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF
; FILE REFERENCE: 0575/58715-A-PCT-US/JPM/ADM/BJA
; CURRENT APPLICATION NUMBER: US/09/932,227
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/04995
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 397
; TYPE: PRT
; ORGANISM: DROSOPHILA MELANOGASTER DOR53
US-09-932-227-8

Query Match 3.6%; Score 7; DB 12; Length 397;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LIPISIS 13
|||||
Db 62 LIPISIS 68

RESULT 43
US-09-738-626-5318
; Sequence 5318, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5318
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5318

Query Match 3.6%; Score 7; DB 10; Length 429;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 80 YTKTCFE 86
|||||
Db 271 YTKTCFE 277

RESULT 44
US-09-934-455-170
; Sequence 170, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddle, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omaira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 170
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-170

Query Match 3.6%; Score 7; DB 11; Length 489;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 TLQSSNS 120
|||||
Db 308 TLQSSNS 314

RESULT 45
US-10-225-068-166
; Sequence 166, Application US/10225068
; Publication No. US20030217383A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Adam, Luc J.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Brown, Pierre E.
; TITLE OF INVENTION: STRESS-RELATED POLYNUCLEOTIDES AND
; FILE REFERENCE: 514442002040
; CURRENT APPLICATION NUMBER: US/10/225,068
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847

PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 246
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 166
LENGTH: 489
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (116)... (179)
OTHER INFORMATION: Conserved domain
US-10-225-068-166

Query Match 3.6%; Score 7; DB 12; Length 489;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 TLOSSNS 120
DB 308 TLOSSNS 314

RESULT 46
US-10-029-386-34057
Sequence 34057, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34057
LENGTH: 495
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007228.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.54
OTHER INFORMATION: SWISSPROT HIT: P52742, EVALUATE 0.00e+00
US-10-029-386-34057

Query Match 3.6%; Score 7; DB 12; Length 495;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SSSLIKH 175
DB 148 SSSLIKH 154

RESULT 47
US-10-369-493-15951
Sequence 15951, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 15951
LENGTH: 537
TYPE: PRT
ORGANISM: Xanthomonas campestris
US-10-369-493-15951

Query Match 3.6%; Score 7; DB 12; Length 537;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 SFFDSND 19
DB 14 SFFDSND 20

RESULT 48
US-10-369-493-16315
Sequence 16315, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 16315
LENGTH: 537
TYPE: PRT
ORGANISM: Xanthomonas campestris
US-10-369-493-16315

Query Match 3.6%; Score 7; DB 12; Length 537;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 SFFDSND 19
DB 14 SFFDSND 20

RESULT 49
US-10-369-493-2798
Sequence 2798, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 2798
LENGTH: 627
TYPE: PRT
ORGANISM: Synechocystis sp.
US-10-369-493-2798

Query Match 3.6%; Score 7; DB 12; Length 627;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 ISLESSG 152
|||||
Db 534 ISLESSG 540

RESULT 50
US-10-047-260-36
Sequence 36, Application US/10047260
Publication No. US20020164706A1
GENERAL INFORMATION:
APPLICANT: Huang, Lisa
APPLICANT: McCluskey, Michael
APPLICANT: Larossa, Robert
TITLE OF INVENTION: HIGH LEVEL PROMOTERS FROM CYANOBACTERIA
FILE REFERENCE: CL1715 US NA
CURRENT APPLICATION NUMBER: US/10/047,260
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/264,925
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Microsoft Office 97
SEQ ID NO 36
LENGTH: 627
TYPE: PRT
ORGANISM: Synechocystis sp. strain PCC6803
US-10-047-260-36

Query Match 3.6%; Score 7; DB 14; Length 627;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 ISLESSG 152
|||||
Db 534 ISLESSG 540

Search completed: January 28, 2004, 18:34:43
Job time : 609 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 28, 2004, 18:12:19 ; Search time 43 Seconds
(without alignments)
431.641 Million cell updates/sec

Title: US-09-881-556A-4
Perfect score: 193
Sequence: 1 HRIFFFLPISISFPDSNDA.....KHAPIKENGHNQESKYSK 193

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 5

Total number of hits satisfying chosen parameters: 11725

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : PIR 76:*

1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	4.1	552	2 S15555	NAD synthase (BC 6
2	7	3.6	120	2 S15057	probable membrane
3	7	3.6	135	2 S76237	hypothetical prote
4	7	3.6	139	2 C70706	hypothetical prote
5	7	3.6	140	2 AB1428	transcription regu
6	7	3.6	140	2 A11801	transcription regu
7	7	3.6	152	2 D49054	T-cell receptor de
8	7	3.6	164	2 T44054	glycoprotein gp82/
9	7	3.6	187	2 P91031	probable fibritai-
10	7	3.6	187	2 G85875	probable major fm
11	7	3.6	203	2 E97022	hypothetical prote
12	7	3.6	212	2 A25856	ubiquitin cholest
13	7	3.6	223	2 JX0322	ubiquitin cholest
14	7	3.6	242	2 A81087	hypothetical prote
15	7	3.6	242	2 D81856	probable integral
16	7	3.6	253	2 S70367	stem cell factor s
17	7	3.6	285	2 S55960	probable membrane
18	7	3.6	287	2 S70366	stem cell factor s
19	7	3.6	287	2 JN0637	stem cell factor p
20	7	3.6	303	2 H75210	probable phosphate
21	7	3.6	477	2 T45722	hypothetical prote
22	7	3.6	602	2 E81398	probable penicilli
23	7	3.6	616	2 T44237	U100 (imported) -
24	7	3.6	627	2 S76378	cell division prot
25	7	3.6	809	2 S32899	ferric-pseudobacti
26	7	3.6	823	2 T29644	hypothetical prote
27	7	3.6	909	2 D87434	phosphoenolpyruvat
28	7	3.6	928	2 T16759	hypothetical prote
29	7	3.6	982	2 T34830	polyketide synthas

30	7	3.6	1068	2 S01519	hypothetical prote
31	7	3.6	1298	2 A64157	hypothetical prote
32	7	3.6	1620	2 S21045	complement protein
33	7	3.6	2938	2 T30249	cell proliferation
34	7	3.6	3169	2 T00296	toxin B - Escheric
35	6	3.1	38	2 E82341	hypothetical prote
36	6	3.1	46	2 P00040	agrd protein - Sta
37	6	3.1	54	2 S40406	blt101 protein - b
38	6	3.1	55	2 H81732	hypothetical prote
39	6	3.1	56	2 T29940	hypothetical prote
40	6	3.1	57	2 T16928	hypothetical prote
41	6	3.1	57	2 T16930	hypothetical prote
42	6	3.1	69	2 E69043	conserved hypochet
43	6	3.1	73	2 F85358	hypothetical prote
44	6	3.1	75	2 C90342	hypothetical prote
45	6	3.1	77	2 T16929	hypothetical prote
46	6	3.1	78	2 C84085	hypothetical prote
47	6	3.1	81	2 T00242	rep protein - Bach
48	6	3.1	87	2 T38941	zinc finger protei
49	6	3.1	90	2 E71527	hypothetical prote
50	6	3.1	90	2 F82571	hypothetical prote
51	6	3.1	94	2 AC1605	comg operon protei
52	6	3.1	94	2 S54639	hypothetical prote
53	6	3.1	98	2 A12161	hypothetical prote
54	6	3.1	99	2 C85707	unknown protein en
55	6	3.1	100	2 A82629	hypothetical prote
56	6	3.1	102	2 T03423	tram protein - Agr
57	6	3.1	102	2 A13243	transcriptional re
58	6	3.1	103	2 J00700	hypothetical 11k p
59	6	3.1	106	2 C65031	biphenyl dioxygena
60	6	3.1	106	2 F91054	phenylpropanate d
61	6	3.1	106	2 A85899	hypothetical prote
62	6	3.1	109	2 T21964	hypothetical prote
63	6	3.1	110	2 C71076	hypothetical prote
64	6	3.1	111	2 S14855	hypothetical prote
65	6	3.1	112	1 WPMG14	14k protein - Cymb
66	6	3.1	113	2 T15934	hypothetical prote
67	6	3.1	114	2 D71048	hypothetical prote
68	6	3.1	115	2 S43487	BCMA complementary
69	6	3.1	115	2 H90475	hypothetical prote
70	6	3.1	118	2 C90914	hypothetical prote
71	6	3.1	119	2 S77748	probable ABC-type
72	6	3.1	120	2 T09392	hypothetical prote
73	6	3.1	121	2 H83768	hypothetical prote
74	6	3.1	122	2 T27074	hypothetical prote
75	6	3.1	124	2 G69821	hypothetical prote
76	6	3.1	124	2 AF0946	hypothetical prote
77	6	3.1	127	2 AB2800	conserved hypochet
78	6	3.1	129	2 F82116	flagellar protein
79	6	3.1	133	2 AH1444	gp59 (Bacteriophag
80	6	3.1	135	2 S52552	chitin variant Th
81	6	3.1	135	2 S52555	chitin variant Th
82	6	3.1	137	2 A97989	hypothetical prote
83	6	3.1	139	2 S60916	probable membrane
84	6	3.1	141	2 A90130	hypothetical prote
85	6	3.1	143	2 A72602	hypothetical prote
86	6	3.1	144	2 E71222	hypothetical prote
87	6	3.1	146	2 S76533	hypothetical prote
88	6	3.1	147	2 D97579	hypothetical prote
89	6	3.1	149	2 C90217	hypothetical prote
90	6	3.1	155	2 S73884	hypothetical prote
91	6	3.1	155	2 C81150	hypothetical prote
92	6	3.1	157	2 D90314	hypothetical prote
93	6	3.1	158	2 E71102	hypothetical prote
94	6	3.1	159	2 A97502	hypothetical prote
95	6	3.1	159	2 AE2720	conserved hypochet
96	6	3.1	162	2 T41667	hypothetical prote
97	6	3.1	165	2 C84934	50S ribosomal prot
98	6	3.1	166	2 C86158	hypothetical prote
99	6	3.1	167	2 S29579	hypothetical prote
100	6	3.1	167	2 G75445	hypothetical prote

ALIGNMENTS

RESULT 1

S15555

NAD synthase (EC 6.3.1.5) adga [validated] - Rhodobacter capsulatus

N:Alternate names: NAD+ synthase (glutamine-hydrolyzing) (EC 6.3.5.1) [misidentification]

C:Species: Rhodobacter capsulatus

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 03-Jun-2002

C/Accession: S15555

R:Willison, J.C.

A:Submitted to the EMBL Data Library, May 1991

A:Description: Essential genes of Rhodobacter capsulatus and Escherichia coli possibly 1

A/Reference number: S15555

A/Accession: S15555

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-552 <WIL>

A/Cross-references: EMBL:X59399; NID:g45985; PIDN:CAA42042.1; PID:g45986

R:Willison, J.C.; Tissock, G.

J. Bacteriol. 176, 3400-3402, 1994

A>Title: The Escherichia coli efg gene and the Rhodobacter capsulatus adga gene code for

A/Reference number: A59276; MUID:9425014; PMID:8195100

A/Contents: annotation; enzyme activity

C/Genetics:

A:Gene: nadB; adga

A/Map position: 38.1 min

C/Function:

A:Description: catalyzes the irreversible reaction of deamido-NAD+, ammonia, and ATP to

A/Note: L-glutamine can act as an ammonia donor, though more slowly

C/Superfamily: Rhodobacter capsulatus NH3-dependent NAD+ synthase

C/Keywords: ligase

Query Match 4.1%; Score 8; DB 2; Length 552;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 FVLNPDGS 31
DB 219 FVLNPDGS 226

RESULT 2

S57057

probable membrane protein YJR038C - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein J1612

C/Species: Saccharomyces cerevisiae

C>Date: 08-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 19-Apr-2002

C/Accession: S57057; S63762

R:Hung, M.E.; Chuat, J.C.; Galibert, F.

submitted to the Protein Sequence Database, September 1995

A/Reference number: S57052

A/Accession: S57057

A/Molecule type: DNA

A/Residues: 1-120 <MAN>

A/Cross-references: EMBL:Z49538; NID:g1015687; PID:g1015688; MIPS:YJR038C

R:Hung, M.E.; Chuat, J.C.; Galibert, F.

Yeast 11, 775-781, 1995

A>Title: Analysis of a 42.5 kb DNA sequence of chromosome X reveals three tRNA genes and

A/Reference number: S63757; MUID:95397595; PMID:7668047

A/Accession: S63762

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-120 <HUA>

A/Cross-references: EMBL:L26344; NID:g1197060; PIDN:AAA8740.1; PID:g1197066

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996

C/Genetics:

A/Cross-references: SGD:S0003799

A/Map position: 10R

A/Note: YJR038C

C/Superfamily: Saccharomyces probable membrane protein YJR038C

C/Keywords: transmembrane protein

F:30-46/Domain: transmembrane #status predicted <TM1>
F:53-69/Domain: transmembrane #status predicted <TM2>

Query Match

3.6%; Score 7; DB 2; Length 120;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SSSLIRK 175
DB 13 SSSLIRK 19

RESULT 3

S76237

hypothetical protein - Synecocystis sp. (strain PCC 6803)

C/Species: Synecocystis sp.

A/Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C/Accession: S76237

R:Kaneko, T.; Sato, S.; Kori, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,

DNA Res. 3, 109-136, 1996

A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

S.

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S76237

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-135 <KAN>

A/Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BA18496.1; PID:d1019221

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 3.6%; Score 7; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 DLATTLQ 116
DB 73 DLATTLQ 79

RESULT 4

C70706

hypothetical protein Rv0760C - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000

C/Accession: C70706

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

; Rastreadam, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: C70706

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-139 <COL>

A/Cross-references: GB:Z80226; GB:AL123456; NID:g3261638; PIDN:CAB02403.1; PID:e266556; I

A/Experimental source: strain H37RV

C/Genetics:

C/Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0760C

Query Match 3.6%; Score 7; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 NPDGSGI 33
DB 48 NPDGSGI 54

```
RESULT 5
AB1428
transcription regulator (Marr family) homolog lmo2827 [imported] - Listeria monocytogenes
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AB1428
R/Glaeer, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Fehl, H.
Science 294, 849-852, 2001
A/Author: Krefc, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; Ma
ok, C.; Schluteter, T.; Simoes, N.; Tietrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A./Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1428
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-140 <GLA>
A/Cross-references: GB:NC_003310; PIDN:CAD01040.1; PID:g16412327; GSPDB:GN00177
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo2827

Query Match 3.6%; Score 7; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ILKNIST 39
DB 131 ILKNIST 137

RESULT 6
AT1801
transcription regulator (Marr family) homolog lin2960 [imported] - Listeria innocua (str
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AT1801
R/Glaeer, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Fehl, H.
Science 294, 849-852, 2001
A/Author: Krefc, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; Ma
ok, C.; Schluteter, T.; Simoes, N.; Tietrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A./Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AT1801
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-140 <GLA>
A/Cross-references: GB:AL592022; PIDN:CA098185.1; PID:g16415500; GSPDB:GN00178
A/Experimental source: strain C1p11262
C/Genetics:
A/Gene: lin2960

Query Match 3.6%; Score 7; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ILKNIST 39
DB 131 ILKNIST 137

RESULT 7
D49054
T-cell receptor delta chain C region - pig (fragment)
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Jul-1999
C/Accession: D49054
R/Thome, A.; Saalmuller, A.; Pfeff, E.
Eur. J. Immunol. 23, 1005-1010, 1993
A/Title: Molecular cloning of porcine T cell receptor alpha, beta, gamma and delta chain
```

```
A/Reference number: A49054; MUID:93238851; PMID:8477796
A/Accession: D49054
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-152 <THO>
A/Cross-references: GB:U21163; NID:g309800; PIDN:AAA65026.1; PID:g309801
A/Note: sequence extracted from NCBI backbone (NCBI:130289)
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: T-cell receptor

Query Match 3.6%; Score 7; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 KISLESS 151
DB 32 KISLESS 38

RESULT 8
T44054
glycoprotein gp82/105 [imported] - human herpesvirus 6 (strain HST) (fragment)
C/Species: human herpesvirus 6
A/Variety: Strain HST
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C/Accession: T44054
R/Isegawa, Y.; Mukai, T.; Nakano, K.; Kagawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kawai
J. Virol. 73, 8053-8063, 1999
A/Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and I
A/Reference number: Z22732; MUID:99412319; PMID:10482554
A/Accession: T44054
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-164 <ISE>
A/Cross-references: EMBL:AB021506; NID:g4995977; PIDN:BA78315.1; PID:g4996082
A/Experimental source: strain HST; pop. variant B
C/Genetics:
A/Note: U98

Query Match 3.6%; Score 7; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 TLOSSNS 120
DB 126 TLOSSNS 132

RESULT 9
F91031
probable fimbrial-like protein [imported] - Escherichia coli (strain O157:H7, substrain f
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002
C/Accession: F91031
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: F91031
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-187 <RAY>
A/Cross-references: GB:BA000007; PIDN:BA036645.1; PID:g13362692; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain R1MD 0509952
C/Genetics:
A/Gene: ECG3222
C/Superfamily: F7-2 fimbrial protein

Query Match 3.6%; Score 7; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 30 GSGILKN 36
 |||||
 Db 60 GSGILKN 66

RESULT 10

G85875

probable major fibribrial subunit 23601 [imported] - Escherichia coli (strain O157:H7, sub
 C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002

C:Accession: G85875

R:Perla, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: G85875

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-187 <STO>

A:Cross-references: GB:AE005174; NID:G12516703; PIDN:AAG57467.1; GSPDB:GN00145; UMGP:Z36

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: Z3601

C:Superfamily: F7-2 fibribrial protein

Query Match 3.6%; Score 7; DB 2; Length 187;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 GSGILKN 36
 |||||
 Db 60 GSGILKN 66

RESULT 11

E97022

hypothetical protein CAC0992 [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: E97022

R:Rolling, J.; Brelton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld

A:Reference number: A96800; MUID:21359325; PMID:21359325

A:Accession: E97022

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-203 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK78968.1; PID:G15023900; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC0992

Query Match 3.6%; Score 7; DB 2; Length 203;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 FEKNEA 91
 |||||
 Db 63 FEKNEA 69

RESULT 12

A25856

ubiquitin thiolesterase (EC 3.1.2.15) PGP9.5 - human

N:Alternate names: gene 9.5 protein; neuron cytosolic protein 9.5; ubiquitin carboxyl-ter

C:Species: Homo sapiens (hmn)

C>Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 10-Dec-1999

C:Accession: A25856; S14307

R:Day, I.N.M.; Thompson, R.J.
 FEBS Lett. 210, 157-160, 1987

A>Title: Molecular cloning of cDNA coding for human PGP 9.5 protein..A novel cytoplasmic
 A:Reference number: A25856; MUID:87080796; PMID:2947814

A:Accession: A25856

A:Molecule type: mRNA

A:Residues: 1-212 <DAY>

A:Cross-references: GB:X04741; NID:G35439; PIDN:CAA28443.1; PID:G35440

A>Note: the authors translated the codon CTC for residue 9 as Val

R:Honore, B.; Rasmussen, H.H.; Vandekerckhove, J.; Celis, J.E.

FEBS Lett. 280, 235-240, 1991

A>Title: Neuronal protein gene product 9.5 (IEF SSP 6104) is expressed in cultured human

A:Reference number: S14307; MUID:91192138; PMID:1849484

A:Accession: S14307

A:Molecule type: protein

A:Residues: 9-14,68-72,73-78,'X',80,95-104,105-112,125-140 <HON>

A:Genetics:

A:Gene: 9.5

C:Function:

A>Description: involved in the hydrolysis of esters and amides at the C-terminal glycine

A>Note: highly expressed in quiescent and proliferating fibroblasts

C:Superfamily: human ubiquitin thiolesterase

C:Keywords: cytosol; thiolester hydrolase

Query Match 3.6%; Score 7; DB 2; Length 212;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 FEKNEA 90
 |||||
 Db 121 FEKNEA 127

RESULT 13

JX0222

ubiquitin thiolesterase (EC 3.1.2.15) PGP9.5 - rat

N:Alternate names: ubiquitin carboxy-terminal esterase; ubiquitin carboxyl-terminal hydro

C:Species: Rattus norvegicus (Norway rat)

C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999

C:Accession: JX0222

R:Kajimoto, Y.; Hashimoto, T.; Shirai, Y.; Nishino, N.; Kuno, T.; Tanaka, C.

J. Biochem. 112, 28-32, 1992

A>Title: cDNA cloning and tissue distribution of a rat ubiquitin carboxyl-terminal hydro

A:Reference number: JX0222; MUID:93054416; PMID:1331034

A:Accession: JX0222

A:Molecule type: mRNA

A:Residues: 1-223 <KAY>

A:Cross-references: DDBJ:DD1109

A:Experimental source: brain

C:Superfamily: human ubiquitin thiolesterase

C:Keywords: thiolester hydrolase

Query Match 3.6%; Score 7; DB 2; Length 223;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 FEKNEA 90
 |||||
 Db 132 FEKNEA 138

RESULT 14

A81087

hypothetical protein NMB1404 [imported] - Neisseria meningitidis (strain MC58 serogroup I

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: A81087

R:Reitelin, H.; Saunders, N.J.; Heideberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Hatt, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappoli, R.; Ver

A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20157555; PMID:10710307

A:Accession: A81087

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <TET>
A:Cross-references: GB:AE002489; GB:AE002098; NID:g7226640; PID:AA41767.1; PID:g722664
A:Experimental source: serogroup B, strain MCS8
C:Genetics:
A:Gene: NMB1404

Query Match 3.6%; Score 7; DB 2; Length 242;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 FLPLPISI 12
|||||
Db 64 FLPLPISI 70

RESULT 15

D81856
Probable integral membrane protein NMA1624 [imported] - Neisseria meningitidis (strain Z
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: D81856
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagerb, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: D81856
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <PAR>
A:Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PID:CA84852.1; PID:g738026
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1624

Query Match 3.6%; Score 7; DB 2; Length 242;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 FLPLPISI 12
|||||
Db 64 FLPLPISI 70

RESULT 16

S70367
stem cell factor short form precursor - quail
C:Species: Coturnix coturnix (quail)
C>Date: 06-Dec-1996 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000
C:Accession: S70367
R:Petitite, J.N.; Kulik, M.J.
Biochim. Biophys. Acta 1307, 149-151, 1996
A>Title: Cloning and characterization of cDNAs encoding two forms of avian stem cell fac
A:Reference number: S70366; MUID:96283808; PMID:8679698
A:Accession: S70367
A:Molecule type: mRNA
A:Residues: 1-253 <PET>
A:Cross-references: EMBL:U43079; NID:g1150877; PID:AA59934.1; PID:g1150878
C:Superfamily: mouse mast cell growth factor
C:Keywords: growth factor; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-253/Product: stem cell factor short form #status predicted <MAT>
F:192-216/Domain: transmembrane #status predicted <TM>

Query Match 3.6%; Score 7; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 178 IKENGHL 184
|||||
Db 127 IKENGHL 133

RESULT 17
S55960
Probable membrane protein YLR404w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein L8084.3
C:Species: Saccharomyces cerevisiae
C>Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 19-Apr-2002
C:Accession: S55960
R:Du, Z.

submitted to the EMBL Data Library, January 1995
A:Description: The sequence of S. cerevisiae cosmid 8084.
A:Reference number: S55944
A:Accession: S55960
A:Molecule type: DNA
A:Residues: 1-285 <DUZ>
A:Cross-references: EMBL:U19729; NID:g625097; PID:g625100; GSPDB:GN00012; MIPS:YLR404w
A:Experimental source: strain S288C (AB972)
C:Genetics:
A:Gene: MIPS:YLR404w
A:Cross-references: SGD:S0004396
A:Map position: 12R

C:Superfamily: Saccharomyces probable membrane protein YLR404w
C:Keywords: transmembrane protein
F:18-34/Domain: transmembrane #status predicted <TM1>
F:249-265/Domain: transmembrane #status predicted <TM2>

Query Match 3.6%; Score 7; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 145 KISLESS 151
|||||
Db 195 KISLESS 201

RESULT 18

S70366
stem cell factor long form precursor - quail
C:Species: Coturnix coturnix (quail)
C>Date: 06-Dec-1996 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000
C:Accession: S70366
R:Petitite, J.N.; Kulik, M.J.
Biochim. Biophys. Acta 1307, 149-151, 1996
A>Title: Cloning and characterization of cDNAs encoding two forms of avian stem cell fact
A:Reference number: S70365; MUID:96283808; PMID:8679698
A:Accession: S70366
A:Molecule type: mRNA
A:Residues: 1-287 <PET>
A:Cross-references: EMBL:U43078; NID:g1150875; PID:AA59933.1; PID:g1150876
C:Superfamily: mouse mast cell growth factor
C:Keywords: growth factor; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-287/Product: stem cell factor long form #status predicted <MAT>
F:226-250/Domain: transmembrane #status predicted <TM>

Query Match 3.6%; Score 7; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 178 IKENGHL 184
|||||
Db 127 IKENGHL 133

RESULT 19

JN0637
stem cell factor precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 21-Jul-2000
C:Accession: JN0637
R:Zhou, J.H.; Ontaki, M.; Sakurai, M.
Gene 127, 269-270, 1993

A>Title: Sequence of a cDNA encoding chicken stem cell factor.
A:Reference number: JN0637; MUID:93273244; PMID:7684722
A:Accession: JN0637
A:Molecule type: mRNA
A:Residues: 1287 <ZHO>
A:Cross-references: GB:D13516; NID:g391648; PIDN:BAA02733.1; PID:g391649
A:Experimental source: brain
C:Superfamily: mouse mast cell growth factor
C:Keywords: growth factor; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-287/Product: stem cell factor #status predicted <MAT>
F:226-248/Domain: transmembrane #status predicted <TM>

Query Match 3.6%; Score 7; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 IKENGHL 184
Db 127 IKENGHL 133

RESULT 20
H75210
probable phosphate regulatory protein PAB0144 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: H75210
R:Anonymous, Genoscope
Submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: H75210
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-303 <KAW>
A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49135.1; PID:g545764
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB0144
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF0472

Query Match 3.6%; Score 7; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 VLNPDS 31
Db 37 VLNPDS 43

RESULT 21
T45722
hypothetical protein FLP2.170 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45722
R:Choisne, N.; Robert, C.; Broctier, P.; Mincker, P.; Catolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z23010
A:Accession: T45722
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <CHO>
A:Cross-references: EMBL:AL132955
A:Experimental source: cultivar Columbia; BAC clone FLP2
C:Genetics:
A:Map position: 3
A>Note: FLP2.170

Query Match 3.6%; Score 7; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 TLQSSNS 120
Db 296 TLQSSNS 302

RESULT 22
E81398
probable penicillin-binding protein Cj0525c [imported] - Campylobacter jejuni (strain NC7
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: E81398
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403; 665-668, 2000
A>Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hpe
A:Reference number: AB1250; MUID:20150912; PMID:10688204
A:Accession: E81398
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-602 <PAR>
A:Cross-references: GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CAB75161.1; PID:g696799;
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: pppB; Cj0525c

Query Match 3.6%; Score 7; DB 2; Length 602;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 SGILKNI 37
Db 193 SGILKNI 199

RESULT 23
T44237
U100 [imported] - human herpesvirus 6 (strain Z29)
C:Species: human herpesvirus 6
A:Variety: strain Z29
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T44237
R:Dominguez, G.; Dambaugh, T.R.; Stamey, F.R.; Dewhurst, S.; Inoue, N.; Pellett, P.E.
J. Virol. 73; 8040-8052, 1999
A>Title: Human herpesvirus 6B genome sequence: coding content and comparison with human
A:Reference number: Z22734; MUID:99412318; PMID:10482553
A:Accession: T44237
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-616 <DOM>
A:Cross-references: EMBL:AF157706; PIDN:RAD49680.1
A:Experimental source: strain Z29; variant B
C:Genetics:
A:Insertions: 129/2; 223/2; 335/2; 361/2; 397/2; 435/2; 489/1; 492/2; 535/2
A:Note: U100

Query Match 3.6%; Score 7; DB 2; Length 616;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 TLQSSNS 120
Db 297 TLQSSNS 303

RESULT 24
S76378
cell division protein ftsH-1 - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein slr0228
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 02-Feb-2001
C:Accession: S76378

R.Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 B:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S76378
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Residues: 1-627 <KAN>
 A:Molecule type: DNA
 A:Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BA10230.1; PID:g100160
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Genetics:
 A:Gene: *ftsH-1*
 C:Superfamily: cell division protein *ftsH*; *ftsH/SEC18/CDG48*-type ATP-binding domain hom
 C:Keywords: ATP; cell division; nucleotide binding; P-loop
 F:1186-395/Domain: *ftsH/SEC18/CDG48*-type ATP-binding domain homology <VATP>
 F:212-219/Region: nucleotide-binding motif A (P-loop)

Query Match 3.6%; Score 7; DB 2; Length 627;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 ISLES5G 152
 DB 534 ISLES5G 540

RESULT 25

S32899
 ferric-pseudobactin receptor *pupB* precursor - *Pseudomonas putida*
 C:Species: *Pseudomonas putida*
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Sep-1999
 C:Accession: S32899
 R:Koeber, M.; van de Vossenbergh, J.; Leong, J.; Weisbeek, P.J.
 M.O1. Microbiol. 8, 591-601, 1993
 A>Title: Identification and characterization of the *pupB* gene encoding an inducible ferr
 A:Reference number: S32899; MUID:93316856; PMID:8392140
 A:Accession: S32899
 A:Molecule type: DNA
 A:Residues: 1-809 <KOS>
 A:Cross-references: GB:X73598; GB:S63481; NID:g403024; PIDN:CA51995.1; PID:g581467
 C:Genetics:
 A:Gene: *pupB*
 A:Start codon: GNG
 C:Superfamily: ferric-pseudobactin receptor; *tonB*-dependent receptor amino-terminal hom
 C:Keywords: membrane protein
 F:1-45/Domain: signal sequence #status predicted <SIG>
 F:46-809/Product: ferric-pseudobactin receptor #status predicted <MAT>
 F:189-319/Domain: *tonB*-dependent receptor amino-terminal homology <TNB>
 F:543-809/Domain: *tonB*-dependent receptor carboxyl-terminal homology <TNC>

Query Match 3.6%; Score 7; DB 2; Length 809;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 LNPDSG 32
 DB 435 LNPDSG 441

RESULT 26

T29644
 hypothetical protein W01A11.3 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T29644
 R:Blanchard, M.; Bradshaw, H.
 submitted to the EMBL Data Library, July 1996
 A:Description: The sequence of *C. elegans* cosmid W01A11.
 A:Reference number: Z20658
 A:Accession: T29644
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-823 <BLA>
 A:Cross-references: EMBL:U64852; PIDN:AA04966.1; GSPDB:GN00023; CESP:W01A11.3
 A:Experimental source: strain Bristol N2; clone W01A11
 C:Genetics:
 A:Gene: *CESP:W01A11.3*
 A:Map position: 5
 A:Introns: 15/3; 87/2; 137/2; 207/1; 258/3; 286/2; 419/3; 594/2; 729/2

Query Match 3.6%; Score 7; DB 2; Length 823;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 SSGENSI 156
 DB 120 SSGENSI 126

RESULT 27

D87434
 phosphoenolpyruvate carboxylase [imported] - *Caulobacter crescentus*
 C:Species: *Caulobacter crescentus*
 C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C:Accession: D87434
 R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Guin, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A>Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: A87249; MUID:21173698; PMID:11255647
 A:Accession: D87434
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-909 <STO>
 A:Cross-references: GB:AE005673; NID:g13422868; PIDN:AA23472.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: *CC1493*
 C:Superfamily: phosphoenolpyruvate carboxylase

Query Match 3.6%; Score 7; DB 2; Length 909;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 GSWLGD 58
 DB 253 GSWLGD 259

RESULT 28

T16759
 hypothetical protein R144.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16759
 R:Favell, T.
 submitted to the EMBL Data Library, March 1995
 A:Description: The sequence of *C. elegans* cosmid R144.
 A:Reference number: Z18571
 A:Accession: T16759
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-928 <FAV>
 A:Cross-references: EMBL:U23515; NID:g746492; PID:g746493; PIDN:AA04545.1; CESP:R144.1
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: *CESP:R144.1*
 A:Introns: 80/3; 125/1; 261/2; 331/3; 350/2; 516/3; 589/1; 615/2; 638/3; 726/1; 865/3; 90

Query Match 3.6%; Score 7; DB 2; Length 928;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NSISTIG 160

Db 848 NSISTIG 854

RESULT 29

T34830
polyketide synthase redx - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 03-Dec-1999
C/Accession: T34830
R/Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, January 1998
A/Reference number: Z21558
A/Accession: T34830
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-982 <OLI>
A/Cross-references: EMBL:AL021530; PIDN:CAA16487.1; GSPDB:GN00070; SCOEDB:SC2E9.19
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: redx; SCOEDB:SC2E9.19

Query Match 3.6%; Score 7; DB 2; Length 982;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 SINDAKSL 23
Db 820 SINDAKSL 826

RESULT 30

S01519
hypothetical protein 1068 - liverwort (Marchantia polymorpha) chloroplast
C/Species: chloroplast Marchantia polymorpha
C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 21-Jul-2000
C/Accession: S01519; A05027
R/Kochi, T.; Shirai, H.; Fukuzawa, H.; Sano, T.; Komano, T.; Umesono, K.; Inokuchi, H.;
J. Mol. Biol. 203, 353-372, 1998
A/Title: Structure and organization of Marchantia polymorpha chloroplast genome. IV. Inv
A/Reference number: S01512; WUID:89068688; PMID:3159437
A/Accession: S01519
A/Molecule type: DNA
A/Residues: 1-1068 <KOH>
A/Cross-references: EMBL:X04455; NID:G11640; PIDN:CAA28143.1; PID:G11732
R/Ohyama, K.; Fukuzawa, H.; Kochi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Shi
Nature 322, 572-574, 1986
A/Title: Chloroplast gene organization deduced from complete sequence of liverwort March
A/Reference number: A38014
A/Contents: annotation; gene organization, sites, features
C/Genetics:
A/Genome: chloroplast
C/Keywords: chloroplast

Query Match 3.6%; Score 7; DB 2; Length 1068;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 LKNISTK 40
Db 908 LKNISTK 914

RESULT 31

A64157
hypothetical protein H10696 - Haemophilus influenzae (strain Rd KW20)
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
C/Accession: A64157
R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995

A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter, A.
A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A/Reference number: A64000; WUID:95350630; PMID:7542800
A/Accession: A64157

A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-1298 <TIGR>
A/Cross-references: GB:U32752; GB:U42023; NID:G1573692; PIDN:ACC22356.1; PID:G1573699; T3
C/Note: best homolog was a hypothetical protein from Escherichia coli
C/Superfamily: ynfN protein

Query Match 3.6%; Score 7; DB 2; Length 1298;
Best Local Similarity 100.0%; Pred. No. 129+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 147 SLESSGE 153
Db 1189 SLESSGE 1195

RESULT 32

S21045
complement protein homolog - inshore hagfish (fragment)
C/Species: Eptatretus burgeri (inshore hagfish)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C/Accession: S21045
R/Ishiguro, H.; Kobayashi, K.; Suzuki, M.; Titani, K.; Tomonaga, S.; Kurosawa, Y.
EMBO J 11, 829-837, 1992
A/Title: Isolation of a hagfish gene that encodes a complement component.
A/Reference number: S21045; WUID:92192016; PMID:1372251
A/Accession: S21045
A/Molecule type: mRNA
A/Residues: 1-1620 <ISH>
A/Cross-references: EMBL:Z11595; NID:G62774; PIDN:CAA7677.1; PID:G62775
C/Superfamily: alpha-2-macroglobulin
C/Keywords: plasma

Query Match 3.6%; Score 7; DB 2; Length 1620;
Best Local Similarity 100.0%; Pred. No. 14+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 VLNPDS 31
Db 346 VLNPDS 352

RESULT 33

T30249
cell proliferation antigen Ki-67 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C/Accession: T30249
R/Staaborg, M.; Gell, K.; Brundell, E.; Hoog, C.
J. Cell Sci. 109, 143-153, 1996
A/Title: The murine Ki-67 cell proliferation antigen accumulates in the nucleolar and het
1 for cell cycle progression.
A/Reference number: Z20787; WUID:96431717; PMID:8834799
A/Accession: T30249
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-2938 <STA>
A/Cross-references: EMBL:X82786; NID:G1177527; PIDN:CAA58026.1; PID:G1177528
A/Experimental source: strain CBA; testis
C/Genetics:
A/Gene: Ki-67
C/Keywords: cell cycle control; nucleus; tandem repeat

Query Match 3.6%; Score 7; DB 2; Length 2938;
Best Local Similarity 100.0%; Pred. No. 2.5+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 145 KISLESS 151
|||||||

DB 2058 KISLESS 2064

RESULT 34

toxin B - Escherichia coli plasmid p0157

C:Species: Escherichia coli

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 17-Nov-2000

C:Accession: T00296; T42195

R:Marking: K.; Ihli, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota,

S.; Shinagawa, H.

DNA Res. 5, 1-9, 1998

A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic

A:Reference number: Z24127; MUID:98290540; PMID:9628576

A:Accession: T00296

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3169 <MA>

A:Cross-references: EMBL:AB011549; NID:94589740; PIDN:BAA31815.1; PID:93337056

A:Experimental source: strain EHEC O157:H7, substrain RMD 0509952

R:Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.

Nucleic Acids Res. 26, 4196-4204, 1998

A:Title: The complete DNA sequence and analysis of the large virulence plasmid of Escher

A:Reference number: Z22068; MUID:98391744; PMID:9722640

A:Accession: T42195

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2236, 'D', 238-1887, 'T', 1888-3169 <BUR>

A:Cross-references: EMBL:AF074613; PIDN:MAC70163.1

A:Experimental source: strain EDJ933; serotype O157:H7

C:Genetics:

A:Gene: toxB

A:Genome: plasmid p0157

A:Note: L7095

C:Superfamily: Escherichia coli probable cytotoxin

C:Keywords: cytotoxin

Query Match 3.6%; Score 7; DB 2; Length 3169;
Best Local Similarity 100.0%; Pred. No. 2; 7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 SGILKNI 37
|||||
DB 314 SGILKNI 320

RESULT 35

hypothetical protein VC0294 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: E82341

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

Charadon, D.; Esmolaev, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragol, I.; Sellers, F.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: E82341

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-38 <HEI>

A:Cross-references: GB:AE004117; GB:AE003852; NID:96554695; PIDN:AA931468.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC0294

A:Map position: 1

Query Match 3.1%; Score 6; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LFLVLP 28

DB 2 LFLVLP 7
|||||

RESULT 36

P00040

agrd protein - Staphylococcus aureus

C:Species: Staphylococcus aureus

C:Date: 07-Sep-1990 #sequence_revision 26-Apr-1996 #text_change 15-Oct-1999

C:Accession: S20796; P00040; S58479

R:Korblum, J.; Projan, S.J.; Kreiswirth, B.N.; Mogazeh, S.L.; Eisen, W.; Rose, H.; Nov

submitted to the EMBL Data Library, March 1990

A:Reference number: S20793

A:Accession: S20796

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-46 <KOR>

A:Cross-references: EMBL:X52543; NID:946505; PIDN:CAA36782.1; PID:946509

R:Janson, L.; Loefdal, S.; Arvidson, S.

Mol. Gen. Genet. 219, 480-485, 1989

A:Title: Identification and nucleotide sequence of the delta-lysin gene, hld, adjacent to

A:Reference number: JQ0386; MUID:90158509; PMID:2622452

A:Accession: P00040

A:Molecule type: DNA

A:Residues: 1-45 <JAN>

R:Novick, R.P.; Projan, S.J.; Korblum, J.; Rose, H.F.; Ji, G.; Kreiswirth, B.; Vandenes

Mol. Gen. Genet. 248, 446-458, 1995

A:Title: The agr P2 operon: an autocatalytic sensory transduction system in Staphylococci

A:Reference number: S58478; MUID:96004766; PMID:7565609

A:Accession: S58479

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-46 <NOV>

A:Cross-references: EMBL:X52543; NID:946505; PIDN:CAA36782.1; PID:946509

Query Match 3.1%; Score 6; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GILKNI 37
|||||
DB 14 GILKNI 19

RESULT 37

S40406

b1c101 protein - barley

C:Species: Hordeum vulgare (barley)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999

C:Accession: S40406

R:Goddard, N.J.; Dunn, M.A.; Zhang, L.; White, A.J.; Jack, P.L.; Hughes, M.A.

Plant Mol. Biol. 23, 871-879, 1993

A:Title: Molecular analysis and spatial expression pattern of a low-temperature-specific

A:Reference number: S40406; MUID:94072734; PMID:8251639

A:Accession: S40406

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-54 <GOD>

A:Cross-references: EMBL:Z25537; NID:9439272; PIDN:CAA80984.1; PID:9439273

C:Superfamily: conserved hypothetical protein b266

Query Match 3.1%; Score 6; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYIP 137
|||||
DB 38 ILGYIP 43

RESULT 38

H81732

hypothetical protein TC0165 [imported] - Chlamydia muridarum (strain N199)

C/Species: Chlamydia muridarum, Chlamydia trachomatis Mopn
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
 C/Accession: H81732
 R/Read: T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A/Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
 A/Reference number: A81500; MUID:20150255; PMID:10684935
 A/Accession: H81732
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-55 <1ET>
 A/Cross-references: GB:AE002284; GB:AE002160; NID:G7190202; PID:AAF39041.1; PID:G7190202
 A/Experimental source: strain Nigg (MOpn)
 C/Genetics:
 A/Gene: TC0165

Query Match 3.1%; Score 6; DB 2; Length 55;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFFELL 8
 Db 22 IFFELL 27

RESULT 39
 T29940
 hypothetical protein T20D4.16 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T29940
 R/Mnx: P.; Graves, T.
 submitted to the EMBL Data Library, November 1996
 A/Description: The sequence of C. elegans cosmid T20D4.
 A/Reference number: Z20712
 A/Accession: T29940
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-56 <MTN>
 A/Cross-references: EMBL:U00029; PIDN:AA837583.1; GSPDB:GN00023; CESP:T20D4.16
 A/Experimental source: strain Bristol N2; clone T20D4
 C/Genetics:
 A/Gene: CESP:T20D4.16
 A/Map position: 5
 A/Introns: 40/3

Query Match 3.1%; Score 6; DB 2; Length 56;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 KEFLGL 49
 Db 48 KEFLGL 53

RESULT 40
 T16928
 hypothetical protein T23F2.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Jun-2000
 C/Accession: T16928
 R/Du: Z.
 submitted to the EMBL Data Library, October 1995
 A/Description: The sequence of C. elegans cosmid T23F2.
 A/Reference number: Z18608
 A/Accession: T16928
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-57 <DUZ>
 A/Cross-references: EMBL:U39649; NID:G1049370; PID:G1049373; PIDN:AAA80385.1; CESP:T23F2
 C/Genetics:
 A/Gene: CESP:T23F2.3

A/Introns: 18/3
 C/Superfamily: conserved hypothetical protein b2666

Query Match 3.1%; Score 6; DB 2; Length 57;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYIP 137
 Db 40 ILGYIP 45

RESULT 41
 T16930
 hypothetical protein T23F2.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Jun-2000
 C/Accession: T16930
 R/Du: Z.
 submitted to the EMBL Data Library, October 1995
 A/Description: The sequence of C. elegans cosmid T23F2.
 A/Reference number: Z18608
 A/Accession: T16930
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-57 <DUZ>
 A/Cross-references: EMBL:U39649; NID:G1049370; PID:G1049375; PIDN:AAA80387.1; CESP:T23F2
 C/Genetics:
 A/Gene: CESP:T23F2.5
 A/Introns: 18/3
 C/Superfamily: conserved hypothetical protein b2666

Query Match 3.1%; Score 6; DB 2; Length 57;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYIP 137
 Db 40 ILGYIP 45

RESULT 42
 E69043
 conserved hypothetical protein MTH138 - Methanobacterium thermoautotrophicum (strain DeJ)
 C/Species: Methanobacterium thermoautotrophicum
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C/Accession: E69043
 R/Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F
 ; Qu, D.; Spadefora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
 A/Reference number: A69000; MUID:98037514; PMID:9371463
 A/Accession: E69043
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-69 <MTN>
 A/Cross-references: GB:AE000986; GB:AE000666; NID:G2622424; PIDN:AA85806.1; PID:G2622424
 A/Experimental source: strain Delta H
 C/Genetics:
 A/Gene: MTH1328
 A/Start codon: TTG

Query Match 3.1%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 DIFLID 63
 Db 59 DIFLID 64

RESULT 43

F85358
 hypothetical protein AT4g30650 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
 C:Accession: F85358
 R:Anonymus, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: AB5001; MUID:2008348; PMID:10617198
 A:Accession: F85358
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-73 <STO>
 A:Cross-references: GB:NC_001268; NID:g7269966; PIDN:CA879783.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4g30650
 A:Map position: 4
 C:Superfamily: conserved hypothetical protein b2666

Query Match
 Best Local Similarity 3.1%; Score 6; DB 2; Length 73;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYIP 137
 |||||
 41 ILGYIP 46

RESULT 44
 C90342
 hypothetical protein S808938 [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: C90342
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
 arrest, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: C90342
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-75 <KUR>
 A:Cross-references: GB:AE006641; NID:g13815050; PIDN:AAK1994.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: S808938

Query Match
 Best Local Similarity 3.1%; Score 6; DB 2; Length 75;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 NKISLE 149
 |||||
 49 NKISLE 54

RESULT 45
 T16929
 hypothetical protein T23F2.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Jun-2000
 C:Accession: T16929
 R:Du, Z.
 submitted to the EMBL Data Library, October 1995
 A:Description: The sequence of C. elegans cosmid T23F2.
 A:Reference number: Z18608
 A:Accession: T16929
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-77 <DUZ>
 A:Cross-references: EMBL:U9649; NID:g1049370; PID:g1049374; PIDN:AAA80386.1; CESP:T23F2
 C:Genetics:

A:Gene: CESP:T23F2.4
 A:introns: 10/3; 38/3
 C:Superfamily: conserved hypothetical protein b2666

Query Match
 Best Local Similarity 3.1%; Score 6; DB 2; Length 77;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYIP 137
 |||||
 60 ILGYIP 65

RESULT 46
 C84085
 hypothetical protein BH3483 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: C84085
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
 Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: AB3650; MUID:20512582; PMID:11058132
 A:Accession: C84085
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-78 <STO>
 A:Cross-references: GB:AP001518; NID:BA000004; NID:g10175792; PIDN:BA807202.1; GSPDB:GN001
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH3483

Query Match
 Best Local Similarity 3.1%; Score 6; DB 2; Length 78;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 ENGLIN 185
 |||||
 17 ENGLIN 22

RESULT 47
 T00242
 rep protein - Escherichia coli plasmid p0157 (fragment)
 N:Alternate names: protein E
 C:Species: Escherichia coli
 C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
 C:Accession: T00242
 R:Makino, K.; Ishii, K.; Yasunaga, T.; Hatpori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota,
 S.; Shingawa, H.
 DNA Res. 5, 1-9, 1998
 A>Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhag
 A:Reference number: Z14127; MUID:98290540; PMID:9628576
 A:Accession: T00242
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-81 <MAK>
 A:Cross-references: EMBL:AB011549; NID:g4589740; PIDN:BA31789.1; PID:g3337030
 A:Experimental source: strain EHEC O157:H7, substrain R1MD 0509952
 C:Genetics:
 A:Gene: rep
 A:Genome: plasmid p0157
 C:Superfamily: 27k replication protein

Query Match
 Best Local Similarity 3.1%; Score 6; DB 2; Length 81;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 PDGSGI 33
 |||||
 5 PDGSGI 10

RESULT 48
I38941
zinc finger protein ZNF154 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 05-Nov-1999
C/Accession: I38941
R/Totomerp, N.; Vissing, H.
Genomics 27, 259-264, 1995
A/Title: Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identified
A/Reference number: A57785; MUID:96044430; PMID:7557990
A/Accession: I38941
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-87 <RES>
A/Cross-references: EMBL:U20648; NID:9676874; PIDN:AAC50257.1; PID:9676875
C/Genetics:
A/Gene: GDB:ZNF154
A/Cross-references: GDB:304989
A/Map position: 19q13.4-19q13.3

Query Match
Best Local Similarity 3.1%; Score 6; DB 2; Length 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Pred. No. 1.2e+02;

QY 170 SSLIKH 175
Db 76 SSLIKH 81

RESULT 49
E71527
hypothetical protein CT330 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C/Species: Chlamydia trachomatis
C/Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C/Accession: E71527
R/Stephens, R.S.; Kaiman, S.; Lamme, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A/Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A/Reference number: A71570; MUID:99000809; PMID:9784136
A/Accession: E71527
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-90 <ARN>
A/Cross-references: GB:AE001306; GB:AE001273; NID:93328748; PIDN:AAC67925.1; PID:9332874
C/Genetics:
A/Experimental source: serotype D, strain UW-3/Cx
A/Gene: CT330

Query Match
Best Local Similarity 3.1%; Score 6; DB 2; Length 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Pred. No. 1.2e+02;

QY 26 LNPDS 31
Db 65 LNPDS 70

RESULT 50
F82571
hypothetical protein XF2321 [imported] - Xylella fastidiosa (strain 9a5c)
C/Species: Xylella fastidiosa
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: F82571
R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A/Reference number: A82515; MUID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: F82571
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-90 <SIM>
A/Cross-references: GB:AE004043; GB:AE003849; NID:99107486; PIDN:AAF85120.1; GSPDB:GN001

A/Experimental source: strain 9a5c
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Brienes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
Submitted to Genbank, June 2000
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
R.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Teshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Ze
A/Reference number: A59328
A/Contents: annotation
C/Genetics:
A/Gene: XF2321

Query Match
Best Local Similarity 3.1%; Score 6; DB 2; Length 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Pred. No. 1.2e+02;

QY 26 LNPDS 31
Db 76 LNPDS 81

Search completed: January 28, 2004, 18:23:32
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 28, 2004, 17:25:20 ; Search time 38 Seconds
(without alignments)
238,846 Million cell updates/sec

Title: US-09-881-556A-4
Perfect score: 193
Sequence: 1 HEIFPFLPISISFPDSNDA.....KHAPIKENGHNNOESKYSK 193

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 5

Total number of hits satisfying chosen parameters: 5726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	8	4.1	552	1	MADE_RHOCA
2	7	3.6	75	1	PSK3_ORISA
3	7	3.6	120	1	YJ08_YEAST
4	7	3.6	163	1	COAD_LACPL
5	7	3.6	223	1	UBL1_HORSE
6	7	3.6	223	1	UBL1_HUMAN
7	7	3.6	223	1	UBL1_MOUSE
8	7	3.6	223	1	UBL1_RAT
9	7	3.6	287	1	SCF_CHICK
10	7	3.6	287	1	SCF_COTUA
11	7	3.6	397	1	O22A_PROME
12	7	3.6	489	1	ZNF1_HUMAN
13	7	3.6	531	1	UD15_RAT
14	7	3.6	627	1	FTH1_SYNY3
15	7	3.6	809	1	PUPB_PSEPU
16	7	3.6	909	1	CAPP_CAUCR
17	7	3.6	1041	1	UNB3_CAREL
18	7	3.6	1068	1	YCF0_MARPO
19	7	3.6	1298	1	YTFN_HABIN
20	7	3.6	1620	1	CO3_EPTBU
21	6	3.1	54	1	L701_HORVU
22	6	3.1	57	1	YCU3_CAREL
23	6	3.1	57	1	YCU3_CAREL
24	6	3.1	57	1	YCU5_CAREL
25	6	3.1	69	1	YD28_MERTH
26	6	3.1	73	1	RC23_ARATY
27	6	3.1	86	1	CXH_BUNMU
28	6	3.1	86	1	NXH1_BUNMU
29	6	3.1	87	1	T154_HUMAN
30	6	3.1	102	1	TRAM_AGRTS
31	6	3.1	103	1	MMOD_METCA
32	6	3.1	106	1	HCAC_ECOLI
33	6	3.1	112	1	Y12K_CMVSI

34	6	3.1	124	1	YHCC_BACSU
35	6	3.1	139	1	YN06_YEAST
36	6	3.1	147	1	RLN5_MYCGA
37	6	3.1	155	1	YNSB_RICPR
38	6	3.1	157	1	YC09_RHIME
39	6	3.1	159	1	YB66_AGRTS
40	6	3.1	159	1	YC09_RHITO
41	6	3.1	161	1	COAD_PROMI
42	6	3.1	165	1	PL10_BUCAI
43	6	3.1	170	1	MS1L_MOUSE
44	6	3.1	171	1	VG56_BPT4
45	6	3.1	175	1	Y773_MERTU
46	6	3.1	181	1	AROL_SALTY
47	6	3.1	185	1	KITH_CBEPU
48	6	3.1	186	1	KITH_CBEPU
49	6	3.1	189	1	MS2L_HUMAN
50	6	3.1	189	1	Y788_MERTU
51	6	3.1	190	1	RUYC_FUSNN
52	6	3.1	194	1	RNFA_BUCBP
53	6	3.1	194	1	YADN_ECOLI
54	6	3.1	195	1	TRIN_SULTO
55	6	3.1	195	1	YL47_ARCFU
56	6	3.1	204	1	YK13_YEAST
57	6	3.1	206	1	THIE_FUSNN
58	6	3.1	207	1	FGFG_HUMAN
59	6	3.1	207	1	FGFG_RAT
60	6	3.1	207	1	RPOA_EUGAN
61	6	3.1	207	1	Y214_MYCGE
62	6	3.1	208	1	Y214_MYCPN
63	6	3.1	208	1	YACL_LEGPN
64	6	3.1	209	1	YDR5_SCHPO
65	6	3.1	210	1	COBH_MERTU
66	6	3.1	218	1	GT26_SCHUA
67	6	3.1	220	1	GT29_PASHE
68	6	3.1	220	1	YADF_ECOLI
69	6	3.1	225	1	CON2_RHISA
70	6	3.1	228	1	LOLD_BUCAI
71	6	3.1	233	1	YCBR_ECOLI
72	6	3.1	239	1	GIDB_CLOAB
73	6	3.1	240	1	YG1P_YEAST
74	6	3.1	240	1	YQ1Z_BACSU
75	6	3.1	242	1	UB1G_PASMU
76	6	3.1	246	1	GRAH_HUMAN
77	6	3.1	250	1	KLKB_HUMAN
78	6	3.1	251	1	REB1_ECOLI
79	6	3.1	252	1	YJG6_YEAST
80	6	3.1	253	1	CUPP_BOMMO
81	6	3.1	257	1	ARGB_BUCAI
82	6	3.1	260	1	YK17_YEAST
83	6	3.1	262	1	CUTB_SCHPO
84	6	3.1	266	1	YJGL_ECO57
85	6	3.1	271	1	VDLC_HELRU
86	6	3.1	275	1	TAGG_BACSU
87	6	3.1	277	1	RAFR_PEPDE
88	6	3.1	282	1	FA9_RAT
89	6	3.1	282	1	SRG3_CAREL
90	6	3.1	284	1	VDLC_HELPY
91	6	3.1	285	1	HMD4_XENLA
92	6	3.1	290	1	TPG3_HUMAN
93	6	3.1	291	1	YJ38_PYROH
94	6	3.1	292	1	BLAN_ENTLC
95	6	3.1	296	1	YD06_SPRMU
96	6	3.1	297	1	YB63_ARCFU
97	6	3.1	300	1	T2R9_MOUSE
98	6	3.1	303	1	FMT_HELPY
99	6	3.1	303	1	HYPB_AZOVI
100	6	3.1	305	1	HYPB_AZOCH

ALIGNMENTS

P54587	baeillus bu
P53542	saccharomyc
O52350	mycoplasma
O92801	ricetia
P58259	rhizobium m
O88974	agrobacteri
O98353	rhizobium l
O88943	proteus mlt
P57148	buchnera ap
O92223	mus musculu
P39362	bacterioph
O58183	metanococc
O89639	salmonella
O05880	chordistoneu
O05879	chordistoneu
O96007	homo sapien
P81231	metanococc
O89980	fusobacteri
O89640	buchnera ap
P37050	escherichia
O97311	sulfolobus
O28135	archaeoglob
P36079	saccharomyc
O81159	fusobacteri
O43320	homo sapien
O54769	rattus norv
O88194	euglena ana
P47456	mycoplasma
P75477	mycoplasma
P37033	legionella
O13740	schizosacch
O58340	metanococc
P08515	schistosoma
P56598	faecicola he
P36857	escherichia
O98819	rhinipcephal
P57383	buchnera ap
P75856	escherichia
O97C44	clostridium
P53323	saccharomyc
P54537	baeillus bu
O98C16	pasteurella
P20718	homo sapien
O90837	homo sapien
P03856	escherichia
P40364	saccharomyc
P42852	bonnyx mort
P57157	buchnera ap
P36039	saccharomyc
P38937	schizosacch
O8816	escherichia
O92K1	helicobacte
P42953	helicobacte
P43465	pediococcus
P16296	rattus norv
P46571	caenorhabdi
O05370	helicobacte
P53775	xenopus lae
O93244	homo sapien
O58759	pyrococcus
P52653	pyrococcus
O88CMT	streplococc
O29398	archaeoglob
O93K14	mus musculu
P56461	helicobacte
P31880	azotobacter
O43949	azotobacter

RESULT 1

MADE_RHOCA
ID MADE_RHOCA STANDARD; PRT; 552 AA.
AC 003638;
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamine-dependent NAD(+) synthetase (EC 6.3.5.1) (NAD(+) synthase
[glutamine-hydrolyzing]).
GN MADE OR ADGA.
OS Rhodospirillum rubrum (Rhodospirillum rubrum).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Rhodospirillum.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33303 / B10;
RX MEDLINE=93159814; Pubmed=8431308;
RA Willison J.C.;
RT "Biochemical genetics revisited: the use of mutants to study carbon
RT and nitrogen metabolism in the photosynthetic bacteria.";
RL FEMS Microbiol. Rev. 10:1-38(1993).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94253014; Pubmed=8195100;
RA Willison J.C., Tissot G.;
RT "The Escherichia coli efg gene and the Rhodospirillum rubrum adga
RT gene code for NH3-dependent NAD synthetase.";
RL J. Bacteriol. 176:3400-3402(1994).
CC -1- FUNCTION: IT NOT KNOWN IF THIS NAD SYNTHASE USES NH(3) OR
CC GLUTAMINE AS THE N DONOR.
CC -1- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + NH(3) + H(2)O = AMP +
CC diphosphate + NAD(+).
CC -1- PATHWAY: NAD biosynthesis.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAD
CC SYNTHETASE FAMILY.
CC -1- SIMILARITY: Contains 1 CN hydrolase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X59399; CAA4042.1; -.
DR PIR; S15555; S15555.
DR HAMAP; MF_00193; fused; 1.
DR InterPro; IPR003694; NAD_synthase.
DR InterPro; IPR003010; Nucleoside/CNhydrolase.
DR Pfam; PF00795; CN hydrolase; 1.
DR Pfam; PF02540; NAD synthase; 1.
DR TIGRPFAMs; TIGR00552; nade; 1.
DR PROSITE; PSS0263; CN_HYDROLASE; 1.
KW Lysase; NAD; ATP-binding.
FT DOMAIN 5 274 CN HYDROLASE.
FT MOD BIND 275 552 LIGASE.
FT MOD BIND 290 297 ATP (BY SIMILARITY).
FT ACT_SITE 292 292 BY SIMILARITY.
SQ SEQUENCE 552 AA; 59706 MW; 5791E1F29DB0308 CRC64;

Query Match 4.1%; Score 8; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 FVLPDGS 31
Db 219 FVLPDGS 226

RESULT 2
PSK3_ORYSA STANDARD; PRT; 75 AA.

AC Q9FRF9;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Phytosulfokine 3 precursor [Contains: Phytosulfokine-alpha (PSK-
DE alpha) (Phytosulfokine-a); Phytosulfokine-beta (PSK-beta)
DE (Phytosulfokine-b)].
GN PSK3.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hejiao J.,
RA Zismann V., Bai G., Bowman C.L., Fujii C.Y., Van Aken S., Craven B.,
RA Utechtack T.R., Kalak H., Feldblum T.V., Quackenbush J., White O.,
RA Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNB0009F04 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF PSK-ALPHA AND PSK-BETA, CHARACTERIZATION, AND SULFATION.
RX MEDLINE=20200296; Pubmed=10734215;
RA Matsubayashi Y., Takagi L., Sakagami Y.;
RT "Phytosulfokine-alpha, a sulfated pentapeptide, stimulates the
RT proliferation of rice cells by means of specific high- and
RT low-affinity binding sites.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:13357-13362(1997).
CC -1- FUNCTION: PROMOTES PLANT CELL DIFFERENTIATION, ORGANOGENESIS AND
CC SOMATIC EMBRYOGENESIS AS WELL AS CELL PROLIFERATION.
CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -1- PTM: SULFATION IS IMPORTANT FOR ACTIVITY AND FOR THE BINDING TO A
CC PUTATIVE MEMBRANE RECEPTOR.
CC -1- PTM: PSK-ALPHA IS PRODUCED BY ENDOPEPTIDASE DIGESTION. PSK-BETA IS
CC PRODUCED FROM PSK-ALPHA BY EXOPEPTIDASE DIGESTION.
CC -1- SIMILARITY: BELONGS TO THE PHYTOSULFOKINE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AC079830; AAC46077.1; -.
DR Gramene; Q9FRF9; -.
KW Growth factor; Differentiation; Signal; Sulfation; Multigene family.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 66 POTENTIAL.
FT PEPTIDE 67 71 PHYTOSULFOKINE-ALPHA.
FT PEPTIDE 67 70 PHYTOSULFOKINE-BETA.
FT PROPEP 72 75 POTENTIAL.
FT MOD RES 67 67 SULFATION.
FT MOD RES 69 69 SULFATION.
SQ SEQUENCE 75 AA; 8175 MW; 2BA585A13152B5D CRC64;

Query Match 3.6%; Score 7; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LPLPIS 13
Db 14 LPLPIS 20

RESULT 3
Y08_YEAST STANDARD; PRT; 120 AA.

DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Hypothetical 13.0 kDa protein in HUL4-GEFI intergenic region.
 GN YJR038C OR J1612.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxId=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C; PubMed=766047;
 RX MEDLINE=9337595; PubMed=766047;
 RA Huang M.-E.; Chuat J.-C.; Galibert F.;
 RT "Analysis of a 42.5 kb DNA sequence of chromosome X reveals three
 RT rRNA genes and 14 new open reading frames including a gene most
 RT probably belonging to the family of ubiquitin-protein ligases.";
 RL Yeast 11:775-781(1995).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L36344; AAA88740.1; -
 CC EMBL: 249538; CAA98565.1; -
 CC PIR: S57057; S57057.
 DR SGI: S0003799; YJR038C.
 DR SGI: S0003799; YJR038C.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 26 46 POTENTIAL.
 FT TRANSMEM 57 77 POTENTIAL.
 SQ SEQUENCE 120 AA: 13020 MW: 07CE6ABCSA6F29C7 CRC64;
 QY 169 SSSSLIK 175
 DB 13 SSSSLIK 19
 QY 169 SSSSLIK 175
 DB 13 SSSSLIK 19
 RESULT 4
 COAD_LACPL STANDARD; PRT; 163 AA.
 ID COAD_LACPL
 AC Q88V8B;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Phosphopantetheine adenylyltransferase (EC 2.7.7.3) (Pantetheine-
 DE phosphate adenylyltransferase) (PPAT) (Dephospho-CoA
 GN COAD OR KDTB OR LP 2133.
 OS Lactobacillus plantarum.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 NCBI_TaxId=1590;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22480296; PubMed=1256566;
 RX Kleebebezem M.; Boekhorst J.; van Kranenburg R.; Molenaar D.;
 RA Kuipers O.P.; Leer R.; Tarchini R.; Peters S.A.; Sandbrink H.M.;
 RA Flier M.W.E.J.; Silekema W.; Klein Lankhorst R.M.; Bron P.A.;
 RA Hoffer S.M.; Nierop Groot M.N.; Kerkhoven R.; De Vries M.; Ursing B.;
 RA De Vos W.M.; Stezen R.J.;
 RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
 RT Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
 CC -1- FUNCTION: Reversibly transfers an adenylyl group from ATP to 4'-
 CC dephosphopantetheine, yielding dephospho-CoA (dPCoA) and
 CC pyrophosphate (By similarity).

CC -1- CATALYTIC ACTIVITY: ATP + pantetheine 4'-phosphate = diphosphate +
 CC 3'-dephospho-CoA.
 CC -1- PATHWAY: Coenzyme A (CoA) biosynthesis; fourth step.
 CC -1- SUBUNIT: Homohexamer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the bacterial coad family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL935258; CAD64494.1; -
 CC DR HAMAP: MF_00151; -1.
 DR HAMAP: MF_00151; -1.
 KW Transferase; Nucleotidyltransferase; Coenzyme A biosynthesis;
 KW Complete proteome.
 SQ SEQUENCE 163 AA: 18072 MW: 439F3D6506DD3B4C CRC64;
 QY 168 FSSSLIK 174
 DB 126 FSSSLIK 132
 QY 168 FSSSLIK 174
 DB 126 FSSSLIK 132
 QY 168 FSSSLIK 174
 DB 126 FSSSLIK 132
 RESULT 5
 UBL1_HORSE STANDARD; PRT; 223 AA.
 ID UBL1_HORSE
 AC O9GM50; O62662;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (UCH-
 DE L1) (Ubiquitin thiolesterase L1) (Neuron cytoplasmic protein 9.5)
 DE (P9.5) (PGP9.5).
 GN UCHL1.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 NCBI_TaxId=9796;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Sato F.; Harigai N.; Iwanaga T.; Hasegawa T.; Ishida N.;
 RT "cDNA cloning of equine ubiquitin C-terminal hydrolase (PGP9.5)." ;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN (12)
 RP SEQUENCE OF 50-188 FROM N.A.
 RC TISSUE=Brain;
 RA Yanase H.; Kitamura H.; Iwanaga T.; Kanehira K.; Okita K.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: UBIQUITIN-PROTEIN HYDROLASE IS INVOLVED BOTH IN THE
 CC PROCESSING OF UBIQUITIN PRECURSORS AND OF UBIQUINATED PROTEINS.
 CC THIS ENZYME IS A THIOL PROTEASE THAT RECOGNIZE AND HYDROLYZE
 CC A PEPTIDE BOND AT THE C-TERMINAL GLYCINE OF UBIQUITIN.
 CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
 CC ubiquitin + a thiol.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C12.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB049188; BAB13757.1; -

DR EMBL; AB013344; BAA28214.1; -
 DR HSSP; P15374; UCH.
 DR MEROPS; C12.001; -
 DR InterPro; IPR001578; UCH_1.
 DR Pfam; PFO1088; peptidase_C12; 1.
 DR PRINTS; PR00707; UBCTHYDLASE.
 DR PRODOM; PD350662; UCH_1; 1.
 DR PROSITE; PS00140; UCH_1; 1.
 KW Ubl conjugation pathway; Hydrolyase; Thiol protease; Multigene family.
 FT ACT_SITE 90 90 BY SIMILARITY.
 FT ACT_SITE 161 161 BY SIMILARITY.
 FT ACT_SITE 176 176 BY SIMILARITY.
 FT DOMAIN 37 54 UBQUITIN BINDING 1 (POTENTIAL).
 FT DOMAIN 170 178 UBQUITIN BINDING 2 (POTENTIAL).
 FT CONFLICT 187 187 T -> A (IN REF. 2).
 SQ SEQUENCE 223 AA; 24918 MM; A32AF7D0A9711A6 CRC64;
 Query Match 3.6%; Score 7; DB 1; Length 223;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 84 CPERNEA 90
 DB 132 CPERNEA 138
 RESULT 6
 UBL1_HUMAN STANDARD; PRT; 223 AA.
 ID UBL1_HUMAN
 AC P09336;
 DT 01-MAR-1989 (Ref. 10, Created)
 DT 01-NOV-1990 (Ref. 16, Last sequence update)
 DT 15-SEP-2003 (Ref. 42, Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolyase isozyme L1 (EC 3.4.19.12) (UCH-L1) (Ubiquitin thiolesterase L1) (Neuron cytoplasmic protein 9.5) (PGP 9.5) (PGP9.5).
 GN UCHL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 12-223 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=87080796; PubMed=2947814;
 RA Day I.N.M., Thompson R.J.;
 RT "Molecular cloning of cDNA coding for human PGP 9.5 protein. A novel cytoplasmic marker for neurones and neuroendocrine cells.";
 RL Biochem. J. 210:157-160(1987).
 RN [2]
 RP SEQUENCE OF 1-15 FROM N.A.
 RX MEDLINE=90303237; PubMed=2163617;
 RA Day I.N.M., Hinks L.J., Thompson R.J.;
 RT "The structure of the human gene encoding protein product 9.5 (PGP9.5), a neuron-specific ubiquitin C-terminal hydrolase.";
 RL Biochem. J. 268:521-524(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=lung, and muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Collins E.A., Grouse L.H., Derge J.G.,
 RA Krausner R.D., Collins F.S., Wagner C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marulita K., Farmer A.A., Rubin G.W., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshitsuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguailano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McKeown P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wolley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Hailton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smillius D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 20-25; 79-91; 106-123 AND 136-151.
 RX MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., van Damme J., Puype M., Gasser B., Cells J.E.,
 RA Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 RN [5]
 RP ACTIVE SITES.
 RX MEDLINE=96234978; PubMed=8639624;
 RA Larsen C.N., Price J.S., Wilkinson K.D.;
 RT "Substrate binding and catalysis by ubiquitin C-terminal hydrolases: RT identification of two active site residues.";
 RL Biochemistry 35:6735-6744(1996).
 RN [6]
 RP VARIANT MET-93.
 RA Leroy E., Boyer R., Auburger G., Leube B., Uim G., Mezey E., Harta G.,
 RA Brownstein M.J., Jónnalagada S., Chernova T., Dehejia A., Lavedan C.,
 RA Gasser T., Steinbach P.J., Wilkinson K.D., Polymeropoulos M.H.;
 RT "The ubiquitin pathway in Parkinson's disease.";
 RL Nature 355:451-452(1998).
 RN [7]
 RP VARIANT MET-93.
 RX PubMed=10454131;
 RA Harhangi B.S., Farrer M.J., Lincoln S., Bonifati V., Meo G.,
 RA De Michele G., Brice A., Durr A., Martinez B., Gasser T., Berzani B.,
 RA Vaughan J.R., Wood N.W., Hardy J., Oostra B.A., Breakefield M.M.;
 RT "The 1163Met mutation in the ubiquitin carboxy-terminal hydrolase-L1 gene is not observed in European cases with familial Parkinson's disease.";
 RL Neurosci. Lett. 270:1-4(1999).
 RN [8]
 RP VARIANT TYR-18.
 RA Lincoln S., Vaughan J., Wood N., Baker M., Adamson J., Gwinn-Hardy K.,
 RA Lynch T., Hardy J., Farrer M.;
 RT "Low frequency of pathogenic mutations in the ubiquitin carboxy-terminal hydrolase gene in familial Parkinson's disease.";
 RL NeuroReport 10:427-429(1999).
 RN [9]
 RP VARIANT TYR-18.
 RA Liu Y., Fallon L., Laethiel H.A., Liu Z., Lansbury P.T. Jr.;
 RT "The UCH-L1 gene encodes two opposing enzymatic activities that affect alpha-synuclein degradation and Parkinson's disease susceptibility.";
 RL Cell 111:209-218(2002).
 CC -1- FUNCTION: UBQUITIN-PROTEIN HYDROLASE IS INVOLVED BOTH IN THE PROCESSING OF UBQUITIN PRECURSORS AND OF UBQUITINATED PROTEIN. THIS ENZYME IS A THIOL PROTEASE THAT RECOGNIZE AND HYDROLYZE A PEPTIDE BOND AT THE C-TERMINAL GLYCINE OF UBQUITIN.
 CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O = ubiquitin + a thiol.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: NEURONS AND CELLS OF THE DIFFUSE NEUROENDOCRINE SYSTEM AND THEIR TUMORS.
 CC -1- POLYMERISM: Variant Met-93 was reported to be a cause of Parkinson's disease, but this has not been proven.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C12.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; X04741; CAA28443.1; -

```

DR EMBL: X17377; CAA35249.1; -.
DR EMBL: BC000332; AAH00332.1; -.
DR EMBL: BC005117; AAH05117.1; -.
DR EMBL: BC006305; AAH06305.1; -.
DR PIR: A25856; A25856.
DR HSSP: P15374; 1UCH.
DR MEROPS: C12.001; -.
DR Aarhus/Chent-2DPAGE; 6123; IEF.
DR Genew; HGNC:12513; UCHL1.
DR MIM; 191342; UCHL1.
DR InterPro; IPR001578; UCH_1.
DR Pfam; PF01088; Peptidase_C12; 1.
DR PRINTS; PR00707; UBCTHYDRLASE.
DR ProDom; PD350662; UCH_1; 1.
DR PROSITE; PS00140; UCH_1; 1.
KM ubl conjugation pathway; Hydrolase; Thiol protease; Multigene family;
KW Polymorphism.
FT ACT_SITE 90 90
FT ACT_SITE 161 161
FT ACT_SITE 176 176
FT DOMAIN 37 54
FT DOMAIN 170 178
FT VARIANT 18 18
FT VARIANT 93 93
FT VARIANT /Frid=VAR_015677.
FT VARIANT I -> M (responsible for an impaired
FT VARIANT catalytic activity).
FT VARIANT /Frid=VAR_015678.
SO SEQUENCE 223 AA; 24824 MW; C9B972AC4D5DA8A CRC64;

Query Match 3.6%; Score 7; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 CFEKNEA 90
DB 132 CFEKNEA 138

RESULT 7
UBLI_MOUSE STANDARD; PRT; 223 AA.
ID UBLI_MOUSE
AC Q9R0P9; Q9R122;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (UCH-
DE L1) (Ubiquitin thiolesterase L1) (Neuron cytoplasmic protein 9.5)
DE (PGP 9.5) (PGP9.5).
DE UCHL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
OX [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=99400551; Pubmed=10471497;
RX Saitoh K., Wang Y.-L., Sun J.G., Yamanishi T., Sakai Y., Kiyosawa H.,
RA Saitoh K., Ichihara N., Wakana S., Kikuchi T., Wada K.;
RT "Intragenic deletion in the gene encoding ubiquitin carboxy-terminal
RT hydrolase in gad mice."
RL Nat. Genet. 23:47-51(1999).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Pituitary;
RA Marzban G., Grillari J., Reisinger E., Hemetsberger T.,
RA Hohenwarter O., Kallinger H.;
RT "Cloning of the mouse homologue of ubiquitin carboxyl-terminal
RT hydrolase PGP9.5."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=CS7BL/6J; TISSUE=Hippocampus;

```

```

RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kanakawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batsalov S., Casavari T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kociba H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.D., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseuk S.,
RA Hasehizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -1- FUNCTION: UBIQUITIN-PROTEIN HYDROLASE IS INVOLVED BOTH IN THE
CC PROCESSING OF UBIQUITIN PRECURSORS AND OF UBIQUITINATED PROTEINS.
CC THIS ENZYME IS A THIOL PROTEASE THAT RECOGNIZE AND HYDROLYZE
CC A PEPTIDE BOND AT THE C-TERMINAL GLYCINE OF UBIQUITIN.
CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in brain and testis.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C12.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB025313; BAA84083.1; -.
DR EMBL: AF172334; AAD51029.1; -.
DR EMBL: AK013729; BAB28976.1; -.
DR HSSP: P15374; 1UCH.
DR MEROPS: C12.001; -.
DR MGD; MGI:103149; Uchl1.
DR InterPro; IPR001578; UCH_1.
DR Pfam; PF01088; Peptidase_C12; 1.
DR PRINTS; PR00707; UBCTHYDRLASE.
DR ProDom; PD350662; UCH_1; 1.
DR PROSITE; PS00140; UCH_1; 1.
KM ubl conjugation pathway; Hydrolase; Thiol protease; Multigene family.
FT ACT_SITE 90 90
FT ACT_SITE 161 161
FT ACT_SITE 176 176
FT DOMAIN 37 54
FT DOMAIN 170 178
FT CONFLICT 149 149 E -> K (IN REF. 2).
SO SEQUENCE 223 AA; 24838 MW; F1402BF7B0C077EA CRC64;

Query Match 3.6%; Score 7; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 CFEKNEA 90
DB 132 CFEKNEA 138

RESULT 8
UBLI_RAT STANDARD; PRT; 223 AA.
ID UBLI_RAT
AC Q00981;
DT 01-APR-1993 (Rel. 25, Created)

```

```

DT 01-APR-1993 (Rel. 25, last sequence update)
DE 28-FEB-2003 (Rel. 41, last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (UCH-
DE L1) (Ubiquitin thiolesterase L1) (Neuron cytoplasmic protein 9.5)
DE (Ref 9.5) (Ref 9.5).
GN UCHL1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93054416; PubMed=1331034;
RA Kajimoto Y., Hashimoto T., Shirai Y., Nishino N., Kuno T.,
RA Tanaka C.;
RT "cDNA cloning and tissue distribution of a rat ubiquitin carboxyl-
RT terminal hydrolase PGP9.5."
RL J. Biochem. 112:128-32(1992).
CC -1- FUNCTION: UBIQUITIN-PROTEIN HYDROLASE IS INVOLVED BOTH IN THE
CC PROCESSING OF UBIQUITIN PRECURSORS AND OF UBIQUITINATED PROTEINS.
CC THIS ENZYME IS A THIOL PROTEASE THAT RECOGNIZES AND HYDROLYZES
CC A PEPTIDE BOND AT THE C-TERMINAL GLYCINE OF UBIQUITIN.
CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C12.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10699; BAA01541.1; -.
DR HSSP; P15374; UCH.
DR MEROPS; C12.001; -.
DR InterPro; IPR001578; UCH 1.
DR Pfam; PF01088; Peptidase_C12; 1.
DR PRINTS; PRO0707; UBCTHYDLASE.
DR PRODOM; PD35062; UCH_1; 1.
DR PROSITE; PS00140; UCH_1; 1.
KW Ub1 conjugation pathway; Hydrolase; Thiol protease; Multigene family.
FT ACT_SITE 90 90 BY SIMILARITY.
FT ACT_SITE 161 161 BY SIMILARITY.
FT ACT_SITE 176 176 BY SIMILARITY.
FT DOMAIN 37 54 UBIQUITIN BINDING 1 (POTENTIAL).
FT DOMAIN 170 178 UBIQUITIN BINDING 2 (POTENTIAL).
SQ SEQUENCE 223 AA; 24782 MW; F1BA04F55B4625E CRC64;

Query Match 3.6%; Score 7; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93273244; PubMed=7684722;
RA Zhou J., Ohtaki M., Sakurai M.;
RT "Sequence of a cDNA encoding chicken stem cell factor."
RL Gene 127:269-270(1993).
CC -1- FUNCTION: Stimulates the proliferation of mast cells. Able to
CC augment the proliferation of both myeloid and lymphoid
CC hematopoietic progenitors in bone marrow culture. Mediates also
CC cell-cell adhesion. Acts synergistically with other cytokines,
CC probably interleukine (By similarity).
CC -1- SUBUNIT: Homodimer, non-covalently linked (Probable).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Also exists as a
CC secreted soluble form (By similarity).
CC -1- DEVELOPMENTAL STAGE: ACTS IN THE EARLY STAGES OF HEMATOPOIESIS.
CC -1- PTM: A soluble form is produced by proteolytic processing of
CC the extracellular domain (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SCF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D13516; BAA02733.1; -.
DR PIR; JN0637; JN0637.
DR InterPro; IPR003452; SCF.
DR Pfam; PF02404; SCF; 1.
KW Growth factor; Glycoprotein; Transmembrane; Signal; Cell adhesion.
FT SIGNAL 1 25
FT CHAIN 26 287 KIT LIGAND.
FT DOMAIN 26 225 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 226 246 POTENTIAL.
FT DOMAIN 247 287 CYTOPLASMIC (POTENTIAL).
FT DISUFID 29 117 BY SIMILARITY.
FT DISUFID 68 167 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 287 AA; 32328 MW; 6A8556AD152578 CRC64;

Query Match 3.6%; Score 7; DB 1; Length 287;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 84 CFEKNEA 90
DB 132 CFEKNEA 138

RESULT 9
SCF_CHICK STANDARD; PRT; 287 AA.
AC 009108;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kit ligand precursor (C-kit ligand) (Stem cell factor) (SCF) (Mast
DE cell growth factor) (MGF).
GN KITLG OR SCF.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

```

```

RESULT 10
SCF_COTUA STANDARD; PRT; 287 AA.
AC 090314; 090315;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kit ligand precursor (C-kit ligand) (Stem cell factor) (SCF) (Mast
DE cell growth factor) (MGF).
GN KITLG OR SCF.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA MEDLINE=96283808; PubMed=8679698;

```

```

RA Petrice J.N., Kulik M.J.;
RT "Cloning and characterization of cDNAs encoding two forms of avian
RL stem cell factor.";
RL Biochim. Biophys. Acta 1307:149-151(1996).
CC -1- FUNCTION: Stimulates the proliferation of mast cells. Able to
CC augment the proliferation of both myeloid and lymphoid
CC hematopoietic progenitors in bone marrow culture. Mediates also
CC cell-cell adhesion. Acts synergistically with other cytokines,
CC probably interleukins (By similarity).
CC -1- SUBUNIT: Homodimer, non-covalently linked (Probable).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoforms 1 and 2).
CC Also exists as a secreted soluble form (isoform 1 only) (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q90314-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q90314-2; Sequence=VSP 006026;
CC -1- PPM: A soluble form is produced by proteolytic processing of
CC isoform 1 in the extracellular domain.
CC -1- SIMILARITY: BELONGS TO THE SCF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
CC DR EMBL; U43078; AAC59933.1; -.
CC DR EMBL; U43079; AAC59934.1; -.
CC DR InterPro; IPR003452; SCF.
CC DR Pfam; PF02404; SCF; 1.
CC KM Growth factor; Glycoprotein; Transmembrane; Signal; Cell adhesion;
CC Alternative splicing.
CC FT SIGNAL 1 25 POTENTIAL.
CC FT CHAIN 26 287 KIT LIGAND.
CC FT DOMAIN 26 225 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 226 246 POTENTIAL.
CC FT DOMAIN 247 287 CYTOPLASMIC (POTENTIAL).
CC FT DISULFID 29 117 BY SIMILARITY.
CC FT DISULFID 68 167 BY SIMILARITY.
CC FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT FT VARSPLIC 179 213 DSRVAVTFTISFPVPVAASSLRNDISGNTSSNSNK -> E
CC (in isoform 2).
CC FT FT /FTID=VSP 006026.
CC SQ SEQUENCE 287 AA; 32455 MW; ABA81AE422A702E CRC64;
CC
CC Query Match 3.6%; Score 7; DB 1; Length 287;
CC Best Local Similarity 100.0%; Pred. No. 18;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 178 IKENGHL 184
CC |||||
CC Db 127 IKENGHL 133
CC
CC RESULT 11
CC O22A_DROME
CC ID O22A_DROME STANDARD; PRT; 397 AA.
CC AC P81909; O9U6X7;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Odorant receptor 22A.
CC OR22A OR OR22A.1 OR DOR22A.1 OR DOR53 OR AN11 OR CG12193.

```

OS *Drosophila melanogaster* (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN RN
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=Berkelley;
RC MEDLINE=99389723; PubMed=10458908;
RX Gao O., Chees A.;
RT "Identification of candidate *Drosophila* olfactory receptors from
RT genomic DNA sequence."
RL Genomics 60:31-39(1999).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RC MEDLINE=20196006; PubMed=10731132;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RX Manameth P.G., Scherer S.E., Li P.W., Hoekstra R.A., Galle R.F.,
RA Manameth P.G., Scherer S.E., Li P.W., Hoekstra R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.C., Mortman U.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Arif U.F., Aggreyani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballow R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Bocchan M.R., Bouck Y., Brokstein F., Broctier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Sheha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
RA Flocker C., Gabriellian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA Goad A.C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mikhlin N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Port V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Strickas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gbhe R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RT Science 287:2185-2195(2000).
RL [3]
RN RN
RP SEQUENCE OF 5-397 FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=Oregon-R; TISSUE=Antenna;
RC MEDLINE=99189757; PubMed=10089887;
RX Voeshall L.B., Amrein H., Morozov P.S., Rzhetsky A., Axel R.;
RA "A spatial map of olfactory receptor expression in the *Drosophila*
RA antenna."
RT Cell 96:725-736(1999).
RL [4]
RN RN
RP IDENTIFICATION, AND TISSUE SPECIFICITY.
RC TISSUE=Antenna;
RC MEDLINE=9916868; PubMed=10069338;
RX Clyne P.J., Warr C.G., Freeman M.R., Lessing D., Kim J., Carlson J.R.,
RT "A novel family of divergent seven-transmembrane proteins: candidate
RT odorant receptors in *Drosophila*."
RL Neuron 22:327-338(1999).
RL -1- FUNCTION: PROBABLE ROLE IN THE ODORANT RESPONSE, BEING AN ODORANT
RECEPTOR.

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN 17-20 SENSORY NEURONS ON THE
CC MEDIAL-PROXIMAL EDGE OF THE ANTENNA.
CC -1- SIMILARITY: BELONGS TO FAMILY DR-OR OF G-PROTEIN COUPLED
CC RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF003586; AAF51364.1; -.
DR EMBL: AF127923; AAD26358.2; -.
DR FlyBase: FBgn00263398; Or22a.
DR InterPro: IPR004117; Tm_6.
DR Pfam: PF02949; Tm_6; 1.
KM Transmembrane; G-protein coupled receptor; Olfaction;
KM Multigene family.
FT DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 50 70 1 (POTENTIAL).
FT DOMAIN 71 86 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 87 107 2 (POTENTIAL).
FT DOMAIN 108 136 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 137 157 3 (POTENTIAL).
FT DOMAIN 158 182 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 183 203 4 (POTENTIAL).
FT DOMAIN 204 263 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 264 280 5 (POTENTIAL).
FT DOMAIN 281 286 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 287 304 6 (POTENTIAL).
FT DOMAIN 305 356 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 357 377 7 (POTENTIAL).
FT DOMAIN 378 397 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 397 425 8 (POTENTIAL).
FT CONFLICT 376 397 D -> SIPLFLNGLPLFLS (IN REF. 1).
FT CONFLICT 376 397 MVKLAFSVTVIKQFNLAERFQ -> VGVNLDISLILHFF
FT CONFLICT 376 397 (IN REF. 1).
SQ SEQUENCE 397 AA; 46753 MW; EBD20EF06E6E883 CRC64;
Query Match 3.6%; Score 7; DB 1; Length 397;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 LPLPISIS 13
DB 62 LPLPISIS 68
RESULT 12
ZN71_HUMAN STANDARD; PRT; 489 AA.
AC GNQ028; Q15919; OSUC09; G9U0D3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Endothelial zinc finger protein induced by tumor necrosis factor alpha
DE (zinc finger protein 71) (ZNF477).
GN ZNF71 OR EZFT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RP TISSUE=Unbilical vein endothelial cells;
RX MEDLINE=21319880; Pubmed=11426589;
RA Metaki C., Murakami T., Umetani M., Wada Y., Ishii M., Tsutsuni S.,
RA Aburatani H., Hamakubo T., Kodama T.,
RA "A novel zinc finger protein mRNA in human umbilical vein endothelial
RA cells is profoundly induced by tumor necrosis factor alpha.";
RT J. Atheroscler. Thromb. 7:97-103(2000).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RA Lamerding J.E., McCready P.M., Skowronski E., Sakeladakis G.,
RA Burkhardt-Schulz K., Gordon L., Dias J., Scott D., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Darganan L.,
RA Ertler A., Christensen M., Georgescu A., Avila J., Artix C.,
RA Andreise T., Amico-Keller G., Cofield J., Duarte S., Lucas S.,
RA Bryce R., Thomas P., Quan G., Krommiller B., Arellano A., Sanders C.,
RA Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.,
RT "Sequence analysis of a 3.6 Mb region in 19q13.4 between D19S891 to
RT ZNF134.";
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=2388257; Pubmed=12477932;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahay J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 151-235 FROM N.A.
RX MEDLINE=92347859; Pubmed=1639391;
RA Aubry M., Martineau C., Zhang F.R., Zahed L., Fiegiewicz D.,
RA Delattre O., Thomas G., de Jong P.T., Julien J.P., Rouleau G.A.;
RT "Cloning of six new genes with zinc finger motifs mapping to short and
RT long arms of human acrocentric chromosome 22 (p and q11.2).";
RN [5]
RP SEQUENCE OF 431-489 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92372070; Pubmed=1505991;
RA Lichter P., Bray P., Ried T., Dawid I.B., Ward D.C.;
RT "Clustering of C2-H2 zinc finger motif sequences within telomeric and
RT fragile site regions of human chromosomes.";
RN [6]
RP Genomics 13:999-1007(1992).
CC -1- FUNCTION: May be involved in transcriptional regulation.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: Highly expressed in placenta, followed by
CC brain, testis, pancreas, heart, small intestine, muscle, uterus,
CC prostate and peripheral blood leukocytes. Not detected in liver,
CC lung, colon, stomach, salivary and thyroid gland.
CC -1- INDUCTION: By tumor necrosis factor alpha (TNFalpha).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: Contains 13 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF269249; AAF8036.1; -.
DR EMBL: AC007228; AAD23608.1; ALT_INIT.

```

```

DR EMBL; BC014280; AAH14280.1; -.
DR EMBL; X60074; CAC16174.1; -.
DR EMBL; M88357; AAA61315.1; -.
DR PIR; A42825; A42825.
DR HSSP; P07248; 1PAA.
DR Genew; HGNC:13141; ZNF71.
DR MIM; 194545; -.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; zf-C2H2; 13.
DR PRINTS; PRO00048; ZINC_FINGER.
DR PRODEM; PD000003; Znf_C2H2; 11.
DR SMART; SM00355; Znf_C2H2; 13.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 13.
KM Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
  Nuclear protein; Repeat.
FT ZN_FING 130 152 C2H2-TYPE.
FT ZN_FING 158 180 C2H2-TYPE.
FT ZN_FING 186 208 C2H2-TYPE.
FT ZN_FING 214 236 C2H2-TYPE.
FT ZN_FING 242 264 C2H2-TYPE.
FT ZN_FING 270 292 C2H2-TYPE.
FT ZN_FING 298 320 C2H2-TYPE.
FT ZN_FING 326 348 C2H2-TYPE.
FT ZN_FING 354 376 C2H2-TYPE.
FT ZN_FING 382 404 C2H2-TYPE.
FT ZN_FING 410 432 C2H2-TYPE.
FT ZN_FING 438 460 C2H2-TYPE.
FT ZN_FING 466 488 C2H2-TYPE.
SQ SEQUENCE 489 AA; 54497 MW; 6B1156D21FF9295C CRC64;

Query Match 3.6%; Score 7; DB 1; Length 489;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SSSLINK 175
Db 142 SSSLINK 148

RESULT 13
ID UNIS_RAT STANDARD; PRT; 531 AA.
AC 064538;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UDP-glucuronosyltransferase 1-5 precursor, microsomal (EC 2.4.1.17)
DE (UDPGLT) (UGT1*5) (UGT1-05) (UGT1.5) (UGT1A5) (B5).
GN UGT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 1-286 FROM N.A.
RC STRAIN=Mistar;
RX MEDLINE=95332265; Pubmed=7608130;
RA Emi Y., Ikushiro S.I., Iyama T.;
RT "Drug-responsive and tissue-specific alternative expression of
RT multiple first exons in rat UDP-glucuronosyltransferase family 1
RT (UGT1) gene complex."
RL J. Biochem. 117:392-399 (1995).
RN [2]
RP SEQUENCE OF 287-531 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90274676; Pubmed=2113380;
RA Sato H., Koike O., Tanabe K., Kashimata S.;
RT "Isolation and sequencing of rat liver albumin UDP-
RT glucuronosyltransferase cDNA: possible alternative splicing of a common
RT primary transcript."
RL Biochem. Biophys. Res. Commun. 169:260-264 (1990).

```

```

CC -1- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS, AND
CC ENDOGENOUS COMPOUNDS.
CC -1- CATALYTIC ACTIVITY: UDP-glucuronate + acceptor = UDP + acceptor
CC beta-D-glucuronoside.
CC -1- SUBCELLULAR LOCATION: Microsomal.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms may be produced. Isoforms have a
CC different N-terminal domain and a common C-terminal domain of
CC 245 residues;
CC Name=1;
CC -1- IsoId=Q64638-1; Sequence=Displayed;
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D38069; BAA07263.1; -.
CC EMBL; M34007; AAA42312.1; ALT TERM.
CC InterPro; IPR002213; UDP_gluco_trans.
CC Pfam; PF00201; UDPGT; 1.
CC PROSITE; PS00375; UDPGT; 1.
DR Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane; Signal;
KM Multigene family; Microsome; Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 531 UDP-GLUCURONOSYLTRANSFERASE 1-5.
FT TRANSMEM 489 505 POTENTIAL.
FT CARBOHYD 116 116 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 131 131 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 431 431 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 531 AA; 59993 MW; 04148C1BACACAC80 CRC64;

Query Match 3.6%; Score 7; DB 1; Length 531;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ILKNIST 39
Db 113 ILKNIST 119

RESULT 14
ID FTH1_SYNY3 STANDARD; PRT; 627 AA.
AC 055700;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division protein ftsH homolog 1 (EC 3.4.24.-).
GN SUR0228.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; Pubmed=8590279;
RA Kaneo T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 644 to 924 of the genome."
RL DNA Res. 2:153-166 (1995).
CC -1- FUNCTION: SEEMS TO ACT AS AN ATP-DEPENDENT ZINC METALLOPEPTIDASE
CC (BY SIMILARITY).
CC -1- COFACTOR: BINDS 1 ZINC ION (POTENTIAL).
CC

```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M41.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64000; BA10230.1; -.
DR PIR: S76378; S76378.
DR MEROPS: M41.005; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003959; AAA_ATPase_centre.
DR InterPro: IPR005960; AAA_sub.
DR InterPro: IPR005936; FESH.
DR InterPro: IPR006642; Peptidase_M41.
DR InterPro: IPR006025; Zn_MTPeptide.
DR Pfam: PF00004; AAA; 1.
DR Pfam: PF01434; Peptidase_M41; 1.
DR SMART: SM00382; AAA; 1.
DR TIGRFAMs: TIGR01241; FESH_fam; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR PROSITE: PS00674; AAA; 1.
KW Cell division; ATP-binding; Transmembrane; Hydrolase; Metalloprotease;
KW Zinc; Multigene family; Complete proteome.
FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 8 28 POTENTIAL.
FT DOMAIN 29 117 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 118 138 POTENTIAL.
FT DOMAIN 139 627 CYTOPLASMIC (POTENTIAL).
FT NE_BIND 212 219 ATP (POTENTIAL).
FT METAL 433 433 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 434 434 BY SIMILARITY.
FT METAL 437 437 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 627 AA; 68496 MW; 4B2C160213CD0517 CRC64;

Query Match 3.6%; Score 7; DB 1; Length 627;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 ISLESSG 152
DB 534 ISLESSG 540

RESULT 15
PUPB_PSEPU STANDARD; PRT; 809 AA.
AC P3047;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ferric-pseudobactin BN7/BN8 receptor precursor.
GN PUPB.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=WCS358;
RX MEDLINE=93316856; PubMed=8392140;
RA Koster M., van de Vossen J., Leong J., Weisbeek P.J.;
RT "Identification and characterization of the pupB gene encoding an
RT inducible ferric-pseudobactin receptor of Pseudomonas putida
RT WCS358."
RL Mol. Microbiol. 8:591-601(1993).
CC -1- FUNCTION: SPECIFIC RECEPTOR FOR THE SIDEROPHORES FERRIC
CC PSEUDOBACTIN BN8 AND FERRIC PSEUDOBACTIN BN7, IRON CHELATING

```

```

CC MOLECULES THAT ALLOW THE ORGANISM TO EXTRACT IRON FROM THE
CC ENVIRONMENT, SPECIALLY UNDER IRON-RESTRICTED CONDITIONS.
CC -1- SUBCELLULAR LOCATION: Outer membrane.
CC -1- INDUCTION: BY IRON LIMITATION IN ADDITION TO THE PRESENCE OF ONE
CC OF THE TWO COGNATE PSEUDOBACTINS BN8 OR BN7.
CC -1- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X73598; CA51995.1; -.
DR PIR: S32899; S32899.
DR InterPro: IPR000531; TonB_boxC.
DR Pfam: PF00593; TonB_dep_Rec; 1.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; FALSE_NEG.
DR PROSITE: PS01156; TONB_DEPENDENT_REC_2; 1.
KW Outer membrane; Iron transport; Transport; TonB box; Signal; Receptor.
FT SIGNAL 1 45 POTENTIAL.
FT CHAIN 46 809 FERRIC-PSEUDOBACTIN BN7/BN8 RECEPTOR.
FT DOMAIN 156 160 POLY-SER.
FT SITE 792 809 TONB C-TERMINAL BOX.
SQ SEQUENCE 809 AA; 88389 MW; 0B339F6E788A86CD CRC64;

Query Match 3.6%; Score 7; DB 1; Length 809;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 LNPDSG 32
DB 435 LNPDSG 441

RESULT 16
CAP_CAUCAU STANDARD; PRT; 909 AA.
AC Q9A871;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxylase (EC 4.1.1.31) (PEPCase) (PEPC).
GN ppc OR CCI493.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Knout H., Shetty J., Berry K.,
RA Uetrichack T., Tran K., Wolf A., Yamanevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: Through the carboxylation of phosphoenolpyruvate (PEP)
CC it forms oxaloacetate, a four-carbon dicarboxylic acid source for
CC the tricarboxylic acid cycle.
CC -1- CATALYTIC ACTIVITY: Phosphate + oxaloacetate = H(2)O +
CC phosphoenolpyruvate + CO(2).
CC -1- PATHWAY: Tricarboxylic acid cycle.
CC -1- SIMILARITY: BELONGS TO THE PEPCASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AE005824; AKK23472.1; -
 CC PIR; D87434; D87434.
 CC HSSP; P00864; 1F1Y.
 CC TIGR; CC1493; -
 CC HAMAP; MF_00595; -; 1.
 CC Interpro; IPR001449; PEPcase.
 CC Pfam; PF00311; PEPcase; 1.
 CC PRINTS; PR00150; PEPCARBYLASE.
 CC PROSITE; PS00781; PEPcase_1; 1.
 CC PROSITE; PS00393; PEPcase_2; 1.
 CC Lyase; Carbon dioxide fixation; Tricarboxylic acid cycle;
 CC Complete proteome.
 CC FT ACT_SITE 151 BY SIMILARITY.
 CC FT SITE 578 BY SIMILARITY.
 CC SQ SEQUENCE 909 AA; 98873 MW; 9058101672477012 CRC64;

Query Match 3.6%; Score 7; DB 1; Length 909;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 52 GSWLGGD 58
 Db 253 GSWLGGD 259

RESULT 17
 UN83 CAEEL STANDARD; PRT; 1041 AA.
 AC Q23064; Q95WB6; (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Nuclear migration protein unc-83 (Uncoordinated protein 83).
 GN UNC-83 OR W01A11.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 CC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C), FUNCTION, CHARACTERIZATION,
 RP AND INTERACTION WITH UNC-84.
 RC STRAIN=Brictol N2;
 RX MEDLINE=21614706; PubMed=11748140;
 RA Starr D.A., Hermann G.J., Malone C.J., Fixsen W., Pileus J.R.,
 RA Horvitz H.R., Han M.;
 RT "unc-83 encodes a novel component of the nuclear envelope and is
 RT essential for proper nuclear migration.";
 RT Development 128:5039-5050 (2001).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Brictol N2;
 RA Blanchard M., Bradshaw H.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Involved in nuclear migration during development and
 CC thereafter. Not required for centrosome attachment to the nucleus.
 CC Participate in nuclear migration via its interaction with unc-84, May
 CC possibly disrupts the interaction between unc-84 and anc-1
 CC and therefore frees the nucleus from the cytoskeleton and allows
 CC the nuclear migration to proceed.
 CC -1- SUBUNIT: Interacts with unc-84.
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (Potential). The
 CC transmembrane domain associates with the nuclear envelope.

CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=a;
 CC IsoId=Q23064-1; Sequence=Displayed;
 CC Name=b;
 CC IsoId=Q23064-2; Sequence=VSP_007083;
 CC Name=c;
 CC IsoId=Q23064-3; Sequence=VSP_007084;
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in migratory nuclei,
 CC first expressed at the nuclear envelope of migrating hyp7 nuclei,
 CC then, at the bean embryonic stage, it is expressed in hyp7 cells,
 CC P cells and intestinal cells. In adults, it is expressed in a
 CC variety of cell-types, including cells around the pharynx and in
 CC the uterus.
 CC -1- DEVELOPMENTAL STAGE: Expressed in embryos and adults.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AF338767; AAL15621.1; -
 CC DR EMBL; U64852; AAB04966.2; -
 CC DR WormPep; W01A11.3; CE31077.
 CC KW Nuclear protein; Transmembrane; Developmental protein; Coiled coil;
 CC Alternative splicing.
 CC TRANSMEM 1005 1024
 CC FT DOMAIN 785 816 ANCHOR FOR TYPE IV MEMBRANE PROTEIN
 CC FT DOMAIN 931 951 COILED COIL (POTENTIAL).
 CC FT VARSPPLIC 1 66 COILED COIL (POTENTIAL).
 CC FT VARSPPLIC 1 301 Missing (in isoform b).
 CC FT VARSPPLIC 1 301 /FTID=VSP_007083.
 CC FT VARSPPLIC 1 301 Missing (in isoform c).
 CC FT VARSPPLIC 1 301 /FTID=VSP_007084.
 CC SQ SEQUENCE 1041 AA; 117821 MW; BB403A9A7C1A01F CRC64;

Query Match 3.6%; Score 7; DB 1; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 150 SSGENSI 156
 Db 338 SSGENSI 344

RESULT 18
 YCF0 MARPO STANDARD; PRT; 1068 AA.
 ID YCF0 MARPO
 AC P12221; Q9T2G3; (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-1989 (Rel. 12, Last annotation update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 127 kDa protein ycf1 (ORF 1068).
 GN YCF1.
 OS Marchantia polymorpha (Liverwort).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
 OC Marchantiopsida; Marchantiales; Marchantiaceae; Marchantia.
 CC NCBI_TaxID=3197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S.,
 RA Umesono K., Shiki Y., Takeuchi M., Chang Z., Aota S., Inokuchi H.,
 RA Ozeki H.;
 RT "Chloroplast gene organization deduced from complete sequence of
 RT liverwort Marchantia polymorpha chloroplast DNA.";
 RL Nature 322:572-574 (1986).
 CC -1- FUNCTION: NOT YET KNOWN.
 CC -1- SIMILARITY: BELONGS TO THE YCF1 FAMILY.


```

CC -1- CAUTION: IN M. POLYMERPHA THIS ORF IS IN TWO PARTS: ORF 1068
CC (N-TERMINAL) AND ORF 464 (C-TERMINAL). IT COULD BE DUE TO A
CC FRAMESHIFT ERROR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04465; CAA28143.1; -
DR PIR: S01519; S0150.
DR HSSP: P40136; 1K90.
DR InterPro: IPR001123; IYSE
DR Chlороplast; Hypothetical protein.
DR SEQUENCE 1068 AA; 127656 MW; A86465124C0E464D CRC64;
SQ
Query Match 3.6%; Score 7; DB 1; Length 1068;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 LKNISTK 40
Db 908 LKNISTK 914
RESULT 19
YTFN HAEIN STANDARD; PRT; 1298 AA.
ID YTFN HAEIN STANDARD; PRT; 1298 AA.
AC O57523;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein HI0696.
GN HI0696.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OC NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McEweney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
RN [2]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae."
RL Electrophoresis 21:411-429(2000).
CC -1- SIMILARITY: STRONG, TO E.COLI YTFN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL: U32752; AAC22356.1; -
CC PIR: A64157; A64157.
CC TIGR: H10696; -
DR Pfam: PF04357; DUF490; 1.
DR Complete proteome.
DR SEQUENCE 1298 AA; 141222 MW; 20905D8ECD9329E9 CRC64;
SQ
Query Match 3.6%; Score 7; DB 1; Length 1298;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 147 SLESGSE 153
Db 1189 SLESGSE 1195
RESULT 20
CO3_EPTBU STANDARD; PRT; 1620 AA.
ID CO3_EPTBU STANDARD; PRT; 1620 AA.
AC P98094;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C3 [Contains: C3A anaphylatoxin] (fragment).
DE Complement C3 [Contains: C3A anaphylatoxin] (fragment).
GN C3.
OS Ectoparasitoid (Inshore hagfish).
OC Eukaryota; Metazoa; Chordata; Hyperotretii; Myxiniiformes;
OC Myxiniidae; Eptatreinae; Eptatreutae.
OC NCBI_TaxID=7764;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92192016; PubMed=1372251;
RA Ishiguro H., Kobayashi K., Suzuki M., Titani K., Tomonaga S.,
RA Kurosawa Y.;
RT "Isolation of a hagfish gene that encodes a complement component."
RL EMBO J. 11:829-837(1992).
CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC COMPLEMENT SYSTEM. AFTER ACTIVATION (C3B), IT CAN BIND COVALENTLY,
CC VIA ITS REACTIVE THIOESTER, TO CELL SURFACE CARBOHYDRATES OR
CC IMMUNE AGGREGATES. CYCLOSTOMATES C3 APPEARS TO REPRESENT THE
CC COMMON ANCESTOR OF MAMMALIAN C3 AND C4, SHOWING SIMILARITIES TO
CC BOTH PROTEINS.
CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z11595; CAA77677.1; -
DR EMBL: Z11596; CAB63257.1; ALT_SEQ.
DR PIR: S21045; S21045.
DR HSSP: P01024; 1C3D.
DR InterPro: IPR002890; A2M_N.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000159; Macroglobulin2.
DR InterPro: IPR001134; Netrin_C.
DR Pfam: PF00207; A2M; 1.
DR Pfam: PF01835; A2M_N; 1.
DR Pfam: PF01821; ANATO; 1.
DR Pfam: PF01759; NTR; 1.
DR ProDom: PD0003264; Anaphylatoxin; 1.
DR SMART: SM00104; ANATO; 1.
DR SMART: SM00643; C345C; 1.
DR PROSITE: PS00477; Alpha 2 MACROGLOBULIN; 1.
DR PROSITE: PS01177; ANAPHYLATOXIN_1; 1.
DR PROSITE: PS01178; ANAPHYLATOXIN_2; 1.
DR Complement pathway; Plasma; Inflammatory response; Glycoprotein;
KW

```

```

KW Thioester bond.
FT NON TER 1
FT CHAIN <1 1620 COMPLEMENT C3.
FT CHAIN <1 633 BETA CHAIN (BY SIMILARITY).
FT CHAIN 634 1336 ALPHA CHAIN (BY SIMILARITY).
FT CHAIN 1343 1620 GAMMA CHAIN (BY SIMILARITY).
FT PEPTIDE 634 714 C3A ANAPHYLATOXIN (BY SIMILARITY).
FT DOMAIN 658 694 ANAPHYLATOXIN-LIKE.
FT DOMAIN 1406 1416 ANAPHYLATOXIN-LIKE.
FT DISULFID 526 780 INTERCHAIN (BY SIMILARITY).
FT DISULFID 593 628 BY SIMILARITY.
FT DISULFID 658 686 BY SIMILARITY.
FT DISULFID 659 693 BY SIMILARITY.
FT DISULFID 672 694 BY SIMILARITY.
FT DISULFID 836 1474 BY SIMILARITY.
FT DISULFID 1062 1114 BY SIMILARITY.
FT DISULFID 1321 1451 BY SIMILARITY.
FT DISULFID 1356 1420 BY SIMILARITY.
FT DISULFID 1467 1472 BY SIMILARITY.
FT DISULFID 1479 1551 BY SIMILARITY.
FT DISULFID 1499 1616 BY SIMILARITY.
FT DISULFID 1598 1607 BY SIMILARITY.
FT CROSSLINK 970 973 Iso-glucanyl cysteine thioester (Cys-Gln)
SQ SEQUENCE 1620 AA; 181747 MW; 21F30FAB72417358 CRC64;
  (By similarity).
Query Match 3.6%; Score 7; DB 1; Length 1620;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 VLNPDS 31
Db 346 VLNPDS 352

RESULT 21
LTO1 HORVU STANDARD; PRT; 54 AA.
ID LTO1 HORVU
AC Q425G9;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Low-temperature induced protein lt101.1 (Blt101) (Salt-stress induced
  hydrophobic peptide ES13).
GN LT101.1 OR LT101 OR ES13.
OS Hordeum vulgare (Barley), and
OC Lophopyrum elongatum (Tail wheatgrass) (Aragopyrum elongatum).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OC NCBI_TaxID=4513, 4588;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=H.vulgare; STRAIN=cv. Igr1; TISSUE=Meristem;
RX MEDLINE=94072734; PubMed=8251639;
RA Goddard N.J., Dunn A.M., Zhang L., White A.J., Jack P.L., Hughes M.A.;
RT "Molecular analysis and spatial expression pattern of a low-
  temperature-specific barley gene, Blt101."
RT Plant Mol. Biol. 23:871-879(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=H.vulgare; STRAIN=cv. Igr1;
RA Brown A.P.C., Dunn A.M., Goddard N.J., Hughes M.A.;
RT "Identification of a novel low-temperature-response element in the
  promoter of the barley gene lt101.1."
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=H.vulgare; TISSUE=Root;
RX MEDLINE=94211914; PubMed=815795;
RA Gulick P.J., Shen W., An H.;
RT "ES13, a stress-induced gene from Lophopyrum elongatum,"
  Plant Physiol. 104:799-800(1994).

```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- INDUCTION: BY COLD SHOCK AND SALT STRESS.
CC -1- SIMILARITY: BELONGS TO THE UPF0057 (PMP3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z25537; CAA80984.1; -
DR EMBL; AJ310994; CAC37081.1; -
DR EMBL; U03966; AAA21847.1; -
DR PIR; S40406; S40406.
DR InterPro; IPR000612; UPF0057.
DR Pfam; PF01679; UPF0057.1.
DR PROSITE; PS01309; UPF0057.1.
KW Transmembrane.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
SQ SEQUENCE 54 AA; 5936 MW; 52521C39468BBB2 CRC64;
Query Match 3.1%; Score 6; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYIP 137
Db 38 ILGYIP 43

RESULT 22
YCU3 CAEEL STANDARD; PRT; 57 AA.
ID YCU3 CAEEL
AC Q22700;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 6.3 kDa protein t23f2.3 in chromosome X.
GN t23f2.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RA Du Z.;
RT "Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
  -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
  -1- SIMILARITY: BELONGS TO THE UPF0057 (PMP3) FAMILY.
  -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@isb-sib.ch).
  CC -----
  DR EMBL; U39649; AAA80385.1; -
  DR PIR; T16928; T16928.
  DR WormPep; T23F2.3; CE04998.
  DR InterPro; IPR000612; UPF0057.
  DR Pfam; PF01679; UPF0057.1.
  DR PROSITE; PS01309; UPF0057.1.
  KW Hypothetical protein; Transmembrane.
  FT TRANSMEM 4 24 POTENTIAL.
  FT TRANSMEM 36 56 POTENTIAL.
  SQ SEQUENCE 57 AA; 6251 MW; 7F8352443E5808AC CRC64;

```

Query Match 3.1%; Score 6; DB 1; Length 57;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYP 137
 Db 40 ILGYP 45

RESULT 23

YCU4_CAEEL STANDARD; PRT; 57 AA.
 AC Q22702;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Hypothetical 6.3 kDa protein T23F2.4 in chromosome X.
 GN T23F2.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Du Z.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE UPF0057 (PMP3) FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U39649; AAA80386.1; ALT_INT.
 DR WormBep; T23F2.4; CE04999.
 DR InterPro; IPR000612; UPF0057.
 DR Pfam; PF01679; UPF0057.1.
 DR PROSITE; PS01309; UPF0057.1.
 KM Hypothetical protein; Transmembrane.
 FT TRANSMEM 3 23 POTENTIAL.
 FT TRANSMEM 36 56
 SQ SEQUENCE 57 AA; 6299 MW; 21364A7B46AF2BAA CRC64;

Query Match 3.1%; Score 6; DB 1; Length 57;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYP 137
 Db 40 ILGYP 45

RESULT 24

YCU5_CAEEL STANDARD; PRT; 57 AA.
 AC Q22702;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Hypothetical 6.3 kDa protein T23F2.5 in chromosome X.
 GN T23F2.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;

RA Du Z.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE UPF0057 (PMP3) FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U39649; AAA80387.1; -.
 DR PIR; T16930; T16930.
 DR WormBep; T23F2.5; CE05000.
 DR InterPro; IPR000612; UPF0057.
 DR Pfam; PF01679; UPF0057.1.
 DR PROSITE; PS01309; UPF0057.1.
 KM Hypothetical protein; Transmembrane.
 FT TRANSMEM 3 23 POTENTIAL.
 FT TRANSMEM 36 56
 SQ SEQUENCE 57 AA; 6275 MW; 2A2654626D3F0490 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 57;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYP 137
 Db 40 ILGYP 45

RESULT 25

YD28_METTH STANDARD; PRT; 69 AA.
 AC Q27383;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical transcriptional regulator MTH1328.
 GN MTH1328.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Sater H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- SIMILARITY: BELONGS TO THE PBSX(XRE) FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE000896; AAB85806.1; -.
 DR PIR; E69043; E69043.

DR HSSP; P16117; 1RPD.
 DR InterPro; IPR001387; HTH_3.
 DR Pfam; PF01381; HTH_3; 1.
 DR SMART; SM00530; HTH_XRE; 1.
 KW Hypothetical protein; Transcription regulation; DNA-binding;
 KM Complete proteome.
 FT DNA BIND 16 35 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 69 AA; 7934 MW; 4504B07C0C56846D CRC64;
 Query Match 3.1%; Score 6; DB 1; Length 69;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 58 DIFILD 63
 Db 59 DIFILD 64
 RESULT 26
 RC23 ARATH STANDARD; PRT; 73 AA.
 AC 09M055;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein Atg30650.
 GN ATG30650.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Scheller C., Wambut R., Murphy G., Volckaert G.,
 RA Pohl T., Dueserthoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Anseuge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,
 RA Kreis M., Delany M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohenseel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbusche F.,
 RA Braeken W., Weltjens I., Voet M., Baetiaens I., Aert R., Defoor E.,
 RA Weltenssger T., Botte G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzner B., Brandt A., Peters S., van Staveren M., Dirke W.,
 RA Moolman P., Klein Lankhorst R., Rose M., Haut J., Koetter P.,
 RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLeay K., Mayes R.,
 RA Pectet A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borikova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fichtmann B., Granderath K., Dauner D., Herzi A.,
 RA Neumann O., Aigritou A., Vitale D., Lignotti R., Piravandi E.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Baryes M., Terol J., Torres A.,
 RA Perez-Perez A., Punelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C.,
 RA Fritsman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegler L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Schemking T., Kallick J., Graves T., Harmon G., Edwards J.,
 RA Latrille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mux P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spiehn J., Ryan E., Andrews S., Geisel C., Layman D.,

RA Du H., Ali J., Berghoff A., Jones K., Dione K., Cotton M., Joshua C.,
 RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
 RA Chen E., Maria M., Martensen R., McCombie W.R.,
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana."
 RL Nature 402:769-777 (1999).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: BELONGS TO THE UPF0057 (GMP) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL; AL161577; CAB79783.1; -
 DR PIR; F85358; F85358.
 DR InterPro; IPR000612; UPF0057.
 DR Pfam; PF01679; UPF0057; 1.
 DR PROSITE; PS01309; UPF0057; 1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 4 24 POTENTIAL.
 FT TRANSMEM 37 57 POTENTIAL.
 SQ SEQUENCE 73 AA; 7905 MW; FB25293D015EF606 CRC64;
 Query Match 3.1%; Score 6; DB 1; Length 73;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 132 ILGYIP 137
 Db 41 ILGYIP 46
 RESULT 27
 CXH BUNNU STANDARD; PRT; 86 AA.
 AC P79688;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cardiotoxin homolog TA-cx-like precursor.
 OS Bungarus multicinctus (Many-banded krait).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosaurs; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Elapidae; Bungarinae; Bungarus.
 OC NCBI_TaxId=8616;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Venom gland;
 RX MEDLINE=97142106; PubMed=8988340;
 RA Chang L.S., Lin J.;
 RT "CDNA sequence analysis of a novel cardiotoxin-like protein from
 RT Taiwan banded krait."
 RL Biochem. Mol. Biol. Int. 40:1271-1276(1996).
 RL -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- SIMILARITY: Belongs to the snake toxin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL; Y10870; CAA71820.1; -

```

DR HSSP; P01430; ICOD.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR Prodom; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Toxin; Cardiotoxin; Multigene family; Signal.
FT SIGNAL 1 21
FT CHAIN 1 21 BY SIMILARITY.
FT DISULFID 22 86 CARDIOTOXIN HOMOLOG TA-CTX-LIKE.
FT DISULFID 24 45 BY SIMILARITY.
FT DISULFID 38 62 BY SIMILARITY.
FT DISULFID 66 78 BY SIMILARITY.
FT DISULFID 79 84 BY SIMILARITY.
SQ SEQUENCE 86 AA; 9745 MW; 7CA91B89562FB97B CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 YTKTCF 85
DB 20 YTKTCF 25

RESULT 28
NXH1_BUNMU STANDARD; PRT; 86 AA.
ID NXH1_BUNMU
AC OYGTIO;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Short neurotoxin homolog NTL1 precursor.
OS Bungarus multicinctus (Many-banded krat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Bungarinae; Bungarus.
OC NCBI_TaxID=8616;
OX
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=99060733; PubMed=9844743;
RA Qian Y., Chunyang F., Yi G., Shengli Y.;
RT "The cDNA cloning and sequence analysis of neurotoxin homologues from
RT Bungarus multicinctus.";
RL Biochem. Mol. Biol. Int. 46:821-828(1998).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: Belongs to the snake toxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ232350; CAA11213.1; -
DR HSSP; P01430; ICOD.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR Prodom; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Toxin; Neurotoxin; Multigene family; Signal.
FT SIGNAL 1 21
FT CHAIN 1 21 BY SIMILARITY.
FT DISULFID 22 86 SHORT NEUROTOXIN HOMOLOG NTL1.
FT DISULFID 24 45 BY SIMILARITY.
FT DISULFID 38 62 BY SIMILARITY.
FT DISULFID 66 78 BY SIMILARITY.
FT DISULFID 79 84 BY SIMILARITY.
SQ SEQUENCE 86 AA; 9776 MW; D9F4E958013BCA CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 YTKTCF 85
DB 20 YTKTCF 25

RESULT 29
ID 2154 HUMAN STANDARD; PRT; 87 AA.
AC Q310C;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 154 (Fragment).
GN ZNF154.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
OX
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Insulinoma;
RX MEDLINE=96044430; PubMed=7557990;
RA Tommerup N., Vissing H.;
RT "Isolation and fine mapping of 16 novel human zinc finger-encoding
RT cDNAs identify putative candidate genes for developmental and
RT malignant disorders.";
RL Genomics 27:259-264(1995).
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U20648; AAC50257.1; -
DR PIR; I38941; I38941.
DR HSSP; P08047; ISP2.
DR Genew; HGNC:12939; ZNF154.
DR MTM; 604085; -
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; Zf-C2H2_3.
DR PRINTS; PR00048; ZNCFINGER.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT ZN_FING 1 29
FT ZN_FING 7 29 C2H2-TYPE.
FT ZN_FING 35 57 C2H2-TYPE.
FT ZN_FING 63 85 C2H2-TYPE.
SQ SEQUENCE 87 AA; 10070 MW; 42B9C4E1AC9A92B3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SSLIKH 175
DB 76 SSLIKH 81

```

```

RESULT 30
TRAM AGRT5
ID TRAM AGRT5 STANDARD; PRT; 102 AA.
AC 04452;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcriptional repressor tram.
GN TRAM OR ATU6131 OR AGR PTI BX139.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Plasmid pTIC58.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OC NCBI_TaxID=176299;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=95113793; PubMed=7814335;
RT Hwang I., Cook D.M., Farrand S.K.;
RT "A new regulatory element modulates homoserine lactone-mediated
RT autoinduction of ri plasmid conjugal transfer.";
RT J. Bacteriol. 177:449-458(1995).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Seubal J.C., Kaul R., Monke D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. Sr.,
RA Chapman P., Clendenning J., Decherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RT Science 294:2317-2323(2001).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Hourlo B., Goldman B.S., Cao Y., Askenazi M., Halling L.,
RA Hourlo B., Gordon J., Vaundin M., Iartchouk O., Epp A., Liu F.,
RA Mollan C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cleo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RT Science 294:2323-2328(2001).
CC -I- FUNCTION: NEGATIVELY REGULATES CONJUGATION AND TRA GENES
CC EXPRESSION BY ANTAGONIZING TRAR/AAI DEPENDENT ACTIVATION. TRAM MAY
CC EITHER BIND OR MODIFY TRAR OR AAI MAKING THEM UNAVAILABLE.
CC ALTERNATIVELY, TRAM MAY BIND TRA PROMOTERS PREVENTING TRAR
CC ACTIVATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF010180; AAC17210.1; -
CC EMBL; AF009431; AAL46367.1; -
CC EMBL; AE007939; AAK91095.1; -
CC PIR; A13243; A13243.
CC PIR; T03423; T03423.
CC Conjugation; Transcription regulation; Repressor; Plasmid;
CC Complete proteome.
CC SEQUENCE 102 AA; 11227 MW; 6FD02DF26C88F291 CRC64;

```

```

Query Match 3.1%; Score 6; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 ILGYIP 137
DB 92 ILGYIP 97

RESULT 31
MOD_METCA
ID MOD_METCA STANDARD; PRT; 103 AA.
AC P22867;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Methane monooxygenase component D.
GN MOD.
OS Methylococcus capsulatus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
OC Methylococcaceae; Methylococcus.
OC NCBI_TaxID=414;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=Bath;
RX MEDLINE=9082694; PubMed=2205538;
RA Stainthorpe A.C., Lees V., Salmood G.P.C., Dalton H., Murrell J.C.;
RT "The methane monooxygenase gene cluster of Methylococcus capsulatus
RT (Bath)."
RT Gene 91:27-34(1990).
RN (2)
RP SEQUENCE OF 1-5, MASS SPECTROMETRY, AND CHARACTERIZATION.
RX STRAIN=Bath;
RX MEDLINE=21838998; PubMed=11709550;
RA Merx M., Lipard S.J.;
RT "Why Orf? Characterization of MOD, a long overlooked component of
RT the soluble methane monooxygenase from Methylococcus capsulatus.";
RT J. Biol. Chem. 277:5858-5865(2002).
CC -I- FUNCTION: Not known.
CC -I- SUBUNIT: The soluble methane monooxygenase (MMO) consists of four
CC components A/MMOH (composed of alpha/MMOX, beta/MMOX and
CC gamma/MMOZ), B/MMOB (mmoB), C/MMOR (mmoC) and D/MMOD (mmoD).
CC -I- MASS SPECTROMETRY: MW=11942.2; METHOD=Electrospray.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M90050; AAF04159.1; -
CC PIR; J00700; J00700.
CC Oxidoreductase; Monooxygenase.
CC SEQUENCE 103 AA; 11942 MW; A91924768A568FB0 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 ISLESS 151
DB 69 ISLESS 74

RESULT 32
HCAC_ECOLI
ID HCAC_ECOLI STANDARD; PRT; 106 AA.
AC P7286;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

Query Match	Best Local Similarity	3.1%	Score 6;	DB 1;	Length 106;
Matches	6;	Conservative	0;	Mismatches	0;
Qy	56 GGDIFI 61				
Db	94 GGDIFI 99				
<p>RESULT 33</p> <p>ID 112K_CMWST STANDARD; PRT; 112 AA.</p> <p>AC Q00479;</p> <p>DT 01-DEC-1992 (Rel. 24, Created)</p> <p>DT 01-DEC-1992 (Rel. 24, Last sequence update)</p> <p>DT 01-DEC-1992 (Rel. 24, Last annotation update)</p> <p>DE Hypothetical 12.4 kDa protein.</p> <p>OS Cymbidium mosaic virus (strain Singapore).</p> <p>OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.</p> <p>OX NCBI_TaxID=31725;</p> <p>RN [1]</p> <p>RA MEDLINE=92256805; PubMed=1581564;</p> <p>RA Neo K.K., Wong S.M., Wu M.;</p> <p>RT "Nucleotide sequences of the two ORFs upstream to the coat protein gene of cymbidium mosaic virus.";</p> <p>RL Plant Mol. Biol. 18:1027-1029(1992).</p> <p>CC -----</p> <p>CC This SWISS-PROT entry is copyright. It is produced through a collaboration</p> <p>CC between the Swiss Institute of Bioinformatics and the EMBL outstation -</p> <p>CC the European Bioinformatics Institute. There are no restrictions on its content</p> <p>CC use by non-profit institutions as long as its content is in no way used for</p> <p>CC modified and this statement is not removed. Usage by and for commercial</p> <p>CC entities requires a license agreement (see http://www.isb-sib.ch/announce/</p> <p>CC or send an email to license@isb-sib.ch).</p> <p>CC -----</p> <p>CC EMBL; Y11070; CAA71950.1; -</p> <p>CC EMBL; AE000340; AAC75593.1; -</p> <p>CC EMBL; D90884; BAA16443.1; -</p> <p>CC EMBL; AE005484; AAG57653.1; -</p> <p>CC EMBL; AP002562; BAB36829.1; -</p> <p>CC EMBL; AE015274; AAN44086.1; -</p> <p>CC PIR; A85899; A85899.</p> <p>CC PIR; G65031; G65031.</p> <p>CC PIR; F91054; F91054.</p> <p>CC HSP; P37332; IROT.</p> <p>CC EcoGene; EG13458; hcaC.</p> <p>CC InterPro; IPR005806; Rieske_dom.</p> <p>CC Pfam; PF00355; Rieske; 1.</p> <p>CC KM Aromatic hydrocarbons catabolism; Electron transport; Metal-binding;</p> <p>CC Iron-sulfur; Iron; 2Fe-2S; Complete proteome.</p> <p>CC METAL 42 42 IRON-SULFUR (2FE-2S) (POTENTIAL).</p> <p>CC METAL 44 44 IRON-SULFUR (2FE-2S) (POTENTIAL).</p> <p>CC METAL 62 62 IRON-SULFUR (2FE-2S) (POTENTIAL).</p> <p>CC METAL 65 65 IRON-SULFUR (2FE-2S) (POTENTIAL).</p> <p>CC SQ SEQUENCE 106 AA; 11329 MW; B83149F4126B72D CRC64;</p>					

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL/X62664; CAA44531.1; -
 DR PIR; S20928; MMWG14.
 DR InterPro; IPR001896; Plant_vir_prot.
 DR Pfam; PF01307; Plant_vir_prot; 1.
 DR ProDom; PD001561; Plant_vir_prot; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 112 AA; 12474 MW; EA9C6E0A7520E1A CRC64;

Query Match 3.1%; Score 6; DB 1; Length 112;
 Best Local Similarity 100.0%; Pred. No. 83;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 KSLFVL 26
 Db 12 KSLFVL 17

RESULT 34
 YHCC_BACSU STANDARD; PRT; 124 AA.
 AC P54587;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein yhcC.
 GN yhcC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NC NCB1_TaxID=1423;
 RX SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA MEDLINE=97124185; PubMed=8966498;
 RA Medak M.A., Terpestra P., Holstappel S., Venema G., Bron S.;
 RA "A 22 kb DNA sequence in the cspB-glpPKD region at 75 degrees on the
 RT Bacillus subtilis chromosome.";
 RL Microbiology 142:3021-3026(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertorello M.G., Bessieres P., Bolotin A., Borcher S.,
 RA Bortosa R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruehl C.V., Calwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denton F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entlang K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giesepi G., Gley B.J., Hage K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holstappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Jorle B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazerevic V.,
 RA Lee S.M., Levine A., Liu H., Maunda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moserl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portetelle D., Potwolik S., Prescott A.M.,
 RA Pressac E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takamori A.,
 RA Takanishi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Tobato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis";
 RL Nature 390:249-256(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL/X6983; CAA65686.1; -
 DR EMBL/229108; CAA12731.1; -
 DR PIR; G69821; G69821.
 DR Subtilisin; BGI1581; yhcC.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 22 POTENTIAL.
 SQ SEQUENCE 124 AA; 13958 MW; F9050B1152A38F3C CRC64;

Query Match 3.1%; Score 6; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 EKNEAH 91
 Db 44 EKNEAH 49

RESULT 35
 YN06_YEAST STANDARD; PRT; 139 AA.
 AC P53842;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hypothetical 15.5 kDa protein in PIK1-POL2 intergenic region.
 GN YN266W OR N0800.
 GN Saccharomyces cerevisiae (Baker's yeast).
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NC NCB1_TaxID=4932;
 RX SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RA MEDLINE=96310631; PubMed=8740425;
 RA Sen-Gupta M., Lyck R., Flegel U., Niedenthal R.K., Hegemann J.H.;
 RT "The sequence of a 24,152 bp segment from the left arm of chromosome
 RT XIV from Saccharomyces cerevisiae between the BNI1 and the POL2
 RT genes";
 RL Yeast 12:505-514(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL/X92494; CAA62322.1; -
 DR EMBL/271542; CAA96173.1; -
 DR PIR; S60916; S60916.
 DR SGD; S0005210; YNL266W.
 KW Hypothetical protein; Transmembrane.
 FT DOMAIN 35 33 POLY-SER.
 FT TRANSMEM 57 77 POTENTIAL.
 FT TRANSMEM 57 77 POTENTIAL.
 SQ SEQUENCE 139 AA; 15510 MW; 308D55609058DF16 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred.No.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 167 GFSSSL 172
 |||||
 Db 95 GFSSSL 100

RESULT 36
 RL15_MYCGA STANDARD; PRT; 147 AA.
 ID RL15_MYCGA
 AC 052350;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 50S ribosomal protein L15.
 GN RPL0 OR RPL15.
 OS Mycoplasma gallisepticum.
 CC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 CC NCBI_TaxId=2096;
 CC [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=AS969var.B;
 RX MEDLINE=20325863; PubMed=10867916;
 RA Skamrov A.V., Gol'dman M.A., Peckitova E.S., Bibilashvili R.S.;
 RT "Determination and analysis of the nucleotide sequence of a segment of
 a Mycoplasma gallisepticum strain AS969 chromosome, containing
 operators S10 and trn23-5."
 RT Mol. Biol. (Mosk) 34:390-396(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA
 (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF036708; AAB95405.1; -
 DR InterPro: IPR005749; L15_bact.
 DR InterPro: IPR001196; Ribosomal_L15.
 DR Pfam: PF01305; Ribosomal_L15; I.
 DR TIGRFAMs: TIGR01071; rplO_bact; 1.
 DR PROSITE: PS00475; RIBOSOMAL_L15; FALSE_NEG.
 KW Ribosomal protein; rRNA-binding.
 SQ SEQUENCE 147 AA; 16017 MW; 54570D2E3AD2B972 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 147;
 Best Local Similarity 100.0%; Pred.No.1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 KINKIS 147
 |||||
 Db 125 KINKIS 130

RESULT 37
 NUSB_RICPR STANDARD; PRT; 155 AA.
 ID NUSB_RICPR
 AC 092E01;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE N utilization substance protein B homolog (NusB protein).
 GN NUSB OR RPI2.
 OS Rickettsia prowazekii.
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 CC Rickettsiaceae; Rickettsiae; Rickettsia.
 CC NCBI_TaxId=782;

RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Anderson S.G.E., Zomrodipour A., Anderson J.O.,
 RA Sichteritz-Ponten T., Alsmark U.C.M., Pedowski R.M., Naeelund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria."
 RT Nature 396:133-140(1998).
 CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTION TERMINATION PROCESS
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUSB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ235270; CAA14629.1; ALT_INIT.
 DR HSSP; P04381; 1EY1.
 DR HAMAP; MF 00073; -; 1.
 DR InterPro: IPR006027; NusB.
 DR Pfam: PF01029; NusB; 1.
 KW Transcription termination; RNA-binding; Complete proteome.
 SQ SEQUENCE 155 AA; 17652 MW; B69F86311680A401 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 155;
 Best Local Similarity 100.0%; Pred.No.1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 140 TINKINK 145
 |||||
 Db 3 TINKINK 8

RESULT 38
 YC09_RHIME STANDARD; PRT; 157 AA.
 ID YC09_RHIME
 AC P58259;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein R01209.
 GN R01209 OR SMC01771.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 CC NCBI_TaxId=382;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Baurt J.,
 RA Boissard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaire V., Masuy D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021."
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -1- SIMILARITY: BELONGS TO THE UPF0168 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC -----
DR EMBL: AL591786; CAC45788.1; -.
DR HAMAP: MF_00440; -. 1.
DR InterPro: IPR005144; ATP.
DR InterPro: IPR003796; DUF193.
DR Pfam: PF03477; ATP-cone; 1.
DR TIGRFAMs: TIGR00244; TIGR00244; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 157 AA; 18004 MW; F7AE2C17C680FFA CRC64;
SQ
Query Match 3.1%; Score 6; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 148 LESSGE 153
DB 95 LESSGE 100

RESULT 39
YB66_AGR75 STANDARD; PRT; 159 AA.
ID 08UG74;
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein Atu1166/AGR_C_2157.
GN ATU1166 OR AGR_C_2157.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monke D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D., Sr.,
RA Chapman P., Clendenen J., Deatherage G., Giller W., Grant C.,
RA Kutayvin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao Y., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Curotto B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaubin M., Iartchouk O., Bpp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Lappae C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
RN [1]
RP SIMILARITY: BELONGS TO THE UPF0168 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE009080; AAL42179.1; -.
CC EMBL: AE008045; AAK86970.1; -.
CC PIR: A97502; A97502.

```

```

DR PIR: AE2720; AE2720.
DR HAMAP: MF_00440; -. 1.
DR InterPro: IPR005144; ATP.
DR InterPro: IPR003796; DUF193.
DR Pfam: PF03477; ATP-cone; 1.
DR TIGRFAMs: TIGR00244; TIGR00244; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 159 AA; 18181 MW; 8AC7CD96BD2705D0 CRC64;
SQ
Query Match 3.1%; Score 6; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 148 LESSGE 153
DB 95 LESSGE 100

RESULT 40
YC09_RHIL0 STANDARD; PRT; 159 AA.
ID YC09_RHIL0
AC 0983B3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein mlr8403.
GN MLR8403.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idehara K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
RN [1]
RP SIMILARITY: BELONGS TO THE UPF0168 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AP003014; BBS54293.1; -.
CC HAMAP: MF_00440; -. 1.
CC InterPro: IPR005144; ATP.
CC InterPro: IPR003796; DUF193.
CC Pfam: PF03477; ATP-cone; 1.
CC TIGRFAMs: TIGR00244; TIGR00244; 1.
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 159 AA; 18050 MW; 4040CDE52BB99D1A CRC64;
SQ
Query Match 3.1%; Score 6; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 148 LESSGE 153
DB 95 LESSGE 100

RESULT 41
COAD_PROM1

```

```

ID COAD_PROMI STANDARD; PRT; 161 AA.
AC Q8RSX4;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Phosphoenolpyruvate adenylyltransferase (EC 2.7.7.3) (Pantetheine-
DE phosphate adenylyltransferase) (PPAT) (Dephospho-CoA
DE pyrophosphorylase).
GN COAD.
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=584;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CECT170;
RA Abitia N., Coderech N., Hita B., Merino S., Izquierdo L., Gavin R.,
RA Tomas J.M., Regue M.;
RT "The inner-core lipopolysaccharide biosynthetic wae gene: function
RT and genetic distribution among some enteric bacteria.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Reversibly transfers an adenylyl group from ATP to 4'-
CC phosphoenolpyruvate, yielding dephospho-CoA (dPCoA) and
CC pyrophosphate (by similarity).
CC -1- CATALYTIC ACTIVITY: ATP + pantetheine 4'-phosphate = diphosphate +
CC 3'-dephospho-CoA.
CC -1- PATHWAY: Coenzyme A (CoA) biosynthesis; fourth step.
CC -1- SUBUNIT: Homohexamer (by similarity).
CC -1- SIMILARITY: Belongs to the bacterial coad family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR HMAP: AY075039; AAL78072.1; -.
DR HAMAP: MF_00151; -.
DR InterPro: IPR004820; Cytidylyltransf.
DR InterPro: IPR004821; Cyt tran. rel.
DR InterPro: IPR001980; LPS biosynth.
DR Pfam: PF01467; CTP transf. 2; 1.
DR TIGRPFAMs: TIGR01510; coad_prev_kctB; 1.
DR TIGRPFAMs: TIGR00125; cyt_tran_rel; 1.
DR TRANSFERASE: Nucleotidyltransferase; Coenzyme A biosynthesis.
KW SEQUENCE 161 AA; 17915 MW; 326F066C94FA5BC7 CRC64;
SQ
Query March 3.1%; Score 6; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 SSSLIK 174
Db 128 SSSLIK 133

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; Pubmed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. ABS."
RL Nature 407:81-86 (2000).
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR HMAP: AP001118; BAB12763.1; -.
DR HAMAP: MF_00362; -.
DR InterPro: IPR001790; Ribosomal_L10.
DR InterPro: IPR002363; Ribosomal_L10eub.
DR Pfam: PF00466; Ribosomal_L10; 1.
DR PROSITE: PS01109; RIBOSOMAL_L10; 1.
KW Ribosomal protein; Complete proteome.
KW SEQUENCE 165 AA; 18246 MW; E3C2E0AF9006B94 CRC64;
SQ
Query March 3.1%; Score 6; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 142 KINKIS 147
Db 14 KINKIS 19

```

```

RESULT 42
RL10_BUCAI STANDARD; PRT; 165 AA.
ID RL10_BUCAI
AC P57148;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10.
GN RPLJ OR BU036.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;

```

```

RESULT 43
MS2L_MOUSE STANDARD; PRT; 170 AA.
ID MS2L_MOUSE
AC Q92223; Q9DCR1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Molybdenum cofactor synthetase protein 2 large subunit (Molybdopterin
DE synthetase large subunit) (MPT synthetase large subunit) (MOC52B).
GN MOC52.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99162191; Pubmed=10053003;
RA Stallmeyer B., Drugeon G., Reiss J., Haenni A.L., Mendel R.R.;
RT "Human molybdopterin synthetase gene: identification of a bicistronic
RT transcript with overlapping reading frames.";
RL Am. J. Hum. Genet. 64:698-705 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischiemann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kiehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Sakurai L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carinini P., de Bonaldo M.F.,
RA Brownstein N.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

```

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.",
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Converts molybdopterin precursor 2 into molybdopterin.
CC This requires the incorporation of two sulfur atoms into precursor
CC 2 to generate a dithiolene group.
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SUBUNIT: Heterodimer of a large and a small subunit (By
CC similarity).
CC MISCELLANEOUS: This protein is produced by a bicistronic gene
CC which also produces the small subunit from an overlapping reading
CC frame.
CC -1- SIMILARITY: BELONGS TO THE MOAE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL, AF091872; MAD14601.1; -
DR EMBL, AK002719; BAB22306.1; -
DR HSSP, P30749; IFM0.
DR MGD; MG1:1336894; Moca2.
DR InterPro; IPR003448; Mb_biosynth_Moae.
DR Pfam; PF02391; Moae; 1. biosynth_Moae.
KW Molybdenum cofactor biosynthesis.
SQ SEQUENCE 170 AA; 18632 MW; 394CA8067A9FA018 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 LEAVSY 73
Db 148 LEAVSY 153

RESULT 44
V656_BPT4 STANDARD; PRT; 171 AA.
ID V656_BPT4
AC P39262;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE dCTP pyrophosphatase (EC 3.6.1.12) (Deoxycytidine-triphosphatase)
DE (dCTPase).
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
NCBI_TaxID=10665;
RN NCBI_TaxID=10665;
RP SEQUENCE FROM N.A.
RX MEDLINE=90255970; PubMed=1692800;
RA Frazler M.W., Mosig G.,
RT "The bacteriophage T4 gene mth whose product inhibits late T4 gene
RL expression in an Escherichia coli rpoH (sigma 32) mutant.",
RT Gene 88:7-14(1990).
RP SEQUENCE FROM N.A.
RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Scidham T., Thomas E.,
RT "Bacteriophage T4 genome analysis.",
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: dCTP + H(2)O = dCMP + diphosphate.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL, M30001; AAB07791.1; -
DR EMBL, AF158101; MAD42499.1; -
DR PIR; T10132; T10132.
KW Hydrolyase.
SQ SEQUENCE 171 AA; 20424 MW; 5BDC643E379D3CE1 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 EIKFEL 47
Db 126 EIKFEL 131

RESULT 45
Y773_METUA STANDARD; PRT; 175 AA.
ID Y773_METUA
AC O58183;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0773 precursor.
GN MJ0773.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
NCBI_TaxID=2190;
RN NCBI_TaxID=2190;
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merriam J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhman J.L., Nguyen D.,
RA Uccireback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Huret M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.,
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL jannaschii.",
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL, U67522; AAB98777.1; -
DR PIR; E64396; E64396.
DR TIGR; M07073;
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 33
FT CHAIN 34 175 HYPOTHETICAL PROTEIN MJ0773.
SQ SEQUENCE 175 AA; 20246 MW; 63EE09A737EEAF11 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 SFEDSN 18
 Db 128 SFEDSN 133

RESULT 46
 AROL_SALTY STANDARD; PRT; 181 AA.
 ID AROL_SALTY STANDARD; PRT; 181 AA.
 AC OXREP9;
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DE Shikimate kinase II (EC 2.7.1.71) (SKII).
 GN AROL OR STM0388 OR STY0421 OR T2476.
 OS Salmonella typhimurium, and
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_Taxid=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=2153494; PubMed=11677609;
 RA McClelland W., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=CT18;
 RX MEDLINE=2153494; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Hague A., Hien T.T., Holroyd S., Jorgels K.,
 RA Krogan A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds W., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=TY2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Lion S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burdard V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -1- CATALYTIC ACTIVITY: ATP + shikimate = ADP + shikimate 3-phosphate.
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
 CC fifth step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: Belongs to the shikimate kinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR HAMAP; MF 00109; -; 1.
 DR InterPro; IPR000623; Shik_kinase.
 DR Pfam; PF01202; SKI; 1.
 DR PRINTS; PR01100; SHIKIMTKINASE.
 DR PROSITE; PS01128; SHIKIMATE_KINASE; 1.
 KW Aromatic amino acid biosynthesis; transferase; kinase; ATP-binding;
 KW Complete proteome.
 FT NP BIND 9 16 ATP (BY SIMILARITY).
 SQ SEQUENCE 181 AA; 19804 MW; F2C4B2C36445C54C CRC64;

Query Match 3.1%; Score 6; DB 1; Length 181;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ALAVS 72
 Db 64 ALAVS 69

RESULT 47
 KITH_CFBPV STANDARD; PRT; 185 AA.
 ID KITH_CFBPV STANDARD; PRT; 185 AA.
 AC Q05880;
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Thymidine kinase (EC 2.7.1.21).
 GN TK.
 OS Choriostoneura fumiferana entomopoxvirus (CFBPV).
 OC Viruses; dsDNA viruses, no RNA stage; Foxviridae; Entomopoxvirinae;
 OC Entomopoxvirus B.
 OX NCBI_Taxid=28322;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93107860; PubMed=1469363;
 RA Lyleva V., Fortin Y., Banville M., Arif B., Richardson C.;
 RT "Comparison of the thymidine kinase genes from three
 RT entomopoxviruses.";
 RL J. Gen. Virol. 73:3235-3240(1992).
 CC -1- CATALYTIC ACTIVITY: ATP + thymidine = ADP + thymidine 5'-
 CC phosphate.
 CC -1- SIMILARITY: BELONGS TO THE THYMIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D10681; BAA01527.1; -.
 DR PIR; J01923; J01923. TK-cell.
 DR InterPro; IPR001267; TK_cell.
 DR Pfam; PF00265; TK; 1.
 DR PROSITE; PS00603; TK_CELULAR_TYPE; 1.
 KW Transferase; kinase; DNA synthesis; ATP-binding.
 FT NP BIND 8 15 ATP (BY SIMILARITY).
 SQ SEQUENCE 185 AA; 21298 MW; 33AD4B3A93F5F5AB CRC64;

Query Match 3.1%; Score 6; DB 1; Length 185;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 FFDSD 19
 Db 88 FFDSD 93

RESULT 48
 KITH_CFBPV STANDARD; PRT; 186 AA.
 ID KITH_CFBPV STANDARD; PRT; 186 AA.
 AC Q05879;

```

DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Thymidine kinase (EC 2.7.1.21).
GN TK.
OS Choristoneura biennis entomopoxvirus (CbEPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
OX NCBI_TaxId=10288;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107860; PubMed=1469363;
RA Lytyn V., Fortin V., Banville M., Arif B., Richardson C.;
RT "Comparison of the thymidine kinase genes from three
RT entomopoxviruses".
RL J. Gen. Virol. 73:3235-3240(1992).
CC -1- CATALYTIC ACTIVITY: ATP + thymidine 5'-
CC phosphate.
CC -1- SIMILARITY: BELONGS TO THE THYMIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL: D10680; BAA01526.1; -.
DR PIR: JQ1922; JQ1922.
DR InterPro: IPR001267; TK_cell1.
DR Pfam: PF00265; TK_1.
DR PROSITE: PS00603; TK_CELLULAR_TYPE; 1.
KM Transferrase; kinase; DNA synthesis; ATP-binding.
FT NP_BIND 8 ATP (By SIMILARITY).
SQ SEQUENCE 186 AA; 21426 MW; C6F63B2F4DAAB9C7 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Mismatches 0; Gaps 0;
Matches 6; Conservative 0; Indels 0;

QY 14 FPDSDND 19
DB 89 FPDSDND 94

RESULT 49
MS2L_HUMAN STANDARD; PRT; 188 AA.
ID MS2L_HUMAN
AC 096007;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Molybdenum cofactor synthase protein 2 large subunit (Molybdopterin
DE synthase large subunit) (MPT synthase large subunit) (MOC52B) (MOCO1-
DE B).
GN MOC52 OR MOCO1 OR MCBPE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RX MEDLINE=93162191; PubMed=10053003;
RA Stealmeier B., Drugeon G., Reiss J., Haenni A.L., Mendel R.R.;
RT "Human molybdopterin synthase gene: identification of a bicistronic
RT transcript with overlapping reading frames."
RL Am. J. Hum. Genet. 64:698-705(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99108152; PubMed=9889283;
RA Sloan J., Kinghorn J.R., Unkles S.E.;

```

```

RT "The two subunits of human molybdopterin synthase: evidence for a
RT bicistronic messenger RNA with overlapping reading frames."
RL Nucleic Acids Res. 27:854-858(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=adrenal gland;
RA Huang C., Huang Q., Wu T., Peng Y., Gu Y., Zhang L., Jiang C., Li Y.,
RA Han Z., Wang Y., Chen Z., Fu G.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP VARIANT MOCOD TYPE B LYS-168.
RX MEDLINE=99162192; PubMed=10053004;
RA Reiss J., Dorche C., Stealmeier B., Mendel R.R., Chen N., Zebot M.T.;
RT "Human molybdopterin synthase gene: genomic structure and mutations in
RT molybdenum cofactor deficiency type B."
RL Am. J. Hum. Genet. 64:706-711(1999).
CC -1- FUNCTION: Converts molybdopterin precursor Z into molybdopterin.
CC This requires the incorporation of two sulfur atoms into precursor
CC Z to generate a dithiolene group.
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SUBUNIT: Heterodimer of a large and a small subunit
CC (Probable).
CC -1- TISSUE SPECIFICITY: Highest levels are found in heart and skeletal
CC muscle. Lower levels are present in brain, kidney and pancreas.
CC Very low levels are found in lung and peripheral blood
CC leukocytes.
CC -1- DISEASE: Defects in MOC52 are the cause of molybdenum cofactor
CC deficiency type B (MOCOD type B) [MIM:252150]; an autosomal
CC recessive disease which leads to the pleiotropic loss of all
CC molybdoenzyme activities and is characterized by severe
CC neurological damage, neonatal seizures and early childhood death.
CC -1- MISCELLANEOUS: This protein is produced by a bicistronic gene
CC which also produces the small subunit from an overlapping reading
CC frame.
CC -1- SIMILARITY: BELONGS TO THE MOAE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL: AF091871; AAD14599.1; -.
DR DR EMBL: AF117815; AAD13297.1; -.
DR DR EMBL: AF155659; AAF67478.1; -.
DR PIR: B59370; B59370.
DR HSSP: P30749; 1FMO.
DR Genew: HGNC:7193; MOC52.
DR MIM: 603708; -.
DR MIM: 252150; -.
DR GO: GO:006777; P-Mo-molybdopterin cofactor biosynthesis; TAS.
DR InterPro: IPR003448; Mb_biosynth_Moae.
DR Pfam: PF02391; Moae; 1.
KM Molybdenum cofactor biosynthesis; Disease mutation.
FT VARIANT 168 E -> K (in MOCOD type B).
SQ SEQUENCE 188 AA; 20944 MW; F405235D585621146 CRC64;

Query Match 3.1%; Score 6; DB 1; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Mismatches 0; Gaps 0;
Matches 6; Conservative 0; Indels 0;

QY 68 LEAVSY 73
DB 148 LEAVSY 153

RESULT 50
Y78B_METUA STANDARD; PRT; 189 AA.
ID Y78B_METUA
AC P81231;

```

```

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0785.1.
GN M0785.1.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_Taxid=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=9637999; PubMed=6688087;
RA But C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Giordex A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Ullrich T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO M.JANNASCHII M0795.1 AND M01249.1.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U67523; AAB98790.1; -.
CC TIGR; M0785.1; -.
CC KM Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 20 40 POTENTIAL.
CC FT TRANSMEM 46 66 POTENTIAL.
CC FT TRANSMEM 100 120 POTENTIAL.
CC FT TRANSMEM 126 146 POTENTIAL.
CC SQ SEQUENCE 189 AA; 21527 MW; 7DB9DB3E8E954EDD CRC64;

```

```

Query Match          3.1%; Score 6; DB 1; length 189;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 LLPISI 12
   |||||
Db 58 LLPISI 63

```

Search completed: January 28, 2004, 18:21:20
 Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 28, 2004, 17:27:09 ; Search time 68 Seconds
(without alignments)
732.414 Million cell updates/sec

Title: US-09-881-556A-4
Perfect score: 193
Sequence: 1 HEIFFFLPISISFPDSNDA.....KAPIKENGHNNQESKYSK 193

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 5

Total number of hits satisfying chosen parameters: 31256

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

SPTRMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rv1virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	4.1	45	16	Q8FCP3
2	8	4.1	5020	16	Q8E9W3
3	7	3.6	88	16	Q8NVL0
4	7	3.6	106	4	Q9P0S2
5	7	3.6	135	10	Q8S8D6
6	7	3.6	135	16	P74399
7	7	3.6	139	16	P71817
8	7	3.6	140	16	Q926T1
9	7	3.6	140	16	Q8Y3K8
10	7	3.6	142	6	Q9BCP3
11	7	3.6	153	5	Q18355
12	7	3.6	164	12	Q9MSZ2
13	7	3.6	177	2	Q8VW80
14	7	3.6	182	5	Q18354
15	7	3.6	187	16	Q8X563
16	7	3.6	203	16	Q97KC7

17	7	3.6	210	16	Q8E9I9	Q8E9I9 shewanella
18	7	3.6	218	17	Q8ZTK9	Q8ZTK9 pyrobaculum
19	7	3.6	223	16	Q8E2P2	Q8E2P2 streptococc
20	7	3.6	223	16	Q8DMU2	Q8DMU2 streptococc
21	7	3.6	242	16	Q9JYW4	Q9JYW4 neisseria m
22	7	3.6	242	16	Q9JTV0	Q9JTV0 neisseria m
23	7	3.6	257	16	Q8E0N9	Q8E0N9 streptococc
24	7	3.6	275	16	Q9CKK9	Q9CKK9 pasteurella
25	7	3.6	278	12	P88930	P88930 kapori's sa
26	7	3.6	285	3	Q06058	Q06058 baccharomye
27	7	3.6	294	16	Q8NMR8	Q8NMR8 corynebacte
28	7	3.6	303	17	Q9V263	Q9V263 pyrococcus
29	7	3.6	308	5	Q9VNB4	Q9VNB4 dirosophila
30	7	3.6	315	4	Q9NUP8	Q9NUP8 homo sapien
31	7	3.6	315	5	Q8I099	Q8I099 dirosophila
32	7	3.6	335	5	Q8IS53	Q8IS53 choristoneu
33	7	3.6	340	16	Q9EWX3	Q9EWX3 streptomyce
34	7	3.6	342	4	Q9BSJ3	Q9BSJ3 homo sapien
35	7	3.6	357	5	Q961T4	Q961T4 dirosophila
36	7	3.6	365	5	Q18356	Q18356 dirosophila
37	7	3.6	374	4	Q8N998	Q8N998 homo sapien
38	7	3.6	376	16	Q8PCW8	Q8PCW8 xanthomonas
39	7	3.6	379	16	Q8PPD7	Q8PPD7 xanthomonas
40	7	3.6	425	2	Q9XDF8	Q9XDF8 streptomyce
41	7	3.6	429	16	Q8NQ07	Q8NQ07 corynebacte
42	7	3.6	446	16	Q8D6L6	Q8D6L6 vibrio vuln
43	7	3.6	450	11	Q9DAI9	Q9DAI9 mus musculu
44	7	3.6	472	2	Q9JRN2	Q9JRN2 actinobacill
45	7	3.6	477	10	Q9SN76	Q9SN76 arabidopsis
46	7	3.6	488	5	Q95UR9	Q95UR9 schistocerc
47	7	3.6	489	10	Q93Z00	Q93Z00 arabidopsis
48	7	3.6	512	5	Q95US0	Q95US0 locusta mlg
49	7	3.6	531	11	Q8VD45	Q8VD45 rattus norv
50	7	3.6	533	5	Q9VRP6	Q9VRP6 dirosophila
51	7	3.6	534	5	Q9NIE3	Q9NIE3 dirosophila
52	7	3.6	577	5	Q18353	Q18353 dirosophila
53	7	3.6	602	16	Q9PHY9	Q9PHY9 campylobact
54	7	3.6	616	12	Q9QJ11	Q9QJ11 human herpe
55	7	3.6	636	5	Q8MP77	Q8MP77 dictyostell
56	7	3.6	637	16	Q8RHA7	Q8RHA7 fusobacteri
57	7	3.6	719	5	Q8I3H7	Q8I3H7 plasmodium
58	7	3.6	730	2	P96090	P96090 thermosaeer
59	7	3.6	754	13	Q9BU16	Q9BU16 cyprinus ca
60	7	3.6	755	10	Q9LNM1	Q9LNM1 arabidopsis
61	7	3.6	784	2	Q9JRO1	Q9JRO1 thermosaeer
62	7	3.6	784	4	Q9GMY0	Q9GMY0 homo sapien
63	7	3.6	830	5	Q95UR8	Q95UR8 schistocerc
64	7	3.6	846	5	Q95US1	Q95US1 locusta mlg
65	7	3.6	903	5	Q95V83	Q95V83 dirosophila
66	7	3.6	903	5	Q9NGX8	Q9NGX8 dirosophila
67	7	3.6	903	5	Q9VZ19	Q9VZ19 dirosophila
68	7	3.6	928	5	Q09997	Q09997 caenorhabdi
69	7	3.6	982	16	Q54142	Q54142 streptomyce
70	7	3.6	1041	5	Q23064	Q23064 caenorhabdi
71	7	3.6	1116	16	Q8PR14	Q8PR14 xanthomonas
72	7	3.6	1116	16	Q8PE49	Q8PE49 xanthomonas
73	7	3.6	1237	16	Q8EPM1	Q8EPM1 shewanella
74	7	3.6	1327	5	Q8I2P3	Q8I2P3 plasmodium
75	7	3.6	2338	11	Q6I769	Q6I769 mus musculu
76	7	3.6	3169	2	Q9ZGR4	Q9ZGR4 escherichia
77	7	3.6	3169	2	Q82916	Q82916 escherichia
78	7	3.6	13414	5	Q8I616	Q8I616 cryptospori
79	7	3.1	22	2	Q93AI2	Q93AI2 thiodicillu
80	6	3.1	35	16	Q8E8T2	Q8E8T2 shewanella
81	6	3.1	37	16	Q8CLAS	Q8CLAS yersinia pe
82	6	3.1	38	16	Q9RV63	Q9RV63 vibrio chol
83	6	3.1	42	11	Q63842	Q63842 mus sp. tra
84	6	3.1	46	2	Q53643	Q53643 stabtylococ
85	6	3.1	54	4	Q14849	Q14849 homo sapien
86	6	3.1	54	10	Q94ID7	Q94ID7 arabidopsis
87	6	3.1	54	10	Q8H001	Q8H001 oryza sativ
88	6	3.1	55	7	Q9MM56	Q9MM56 astatoctilap
89	6	3.1	55	7	Q9MM73	Q9MM73 alcolapia a

90 6 3.1 55 16 Q9PUD9
91 6 3.1 56 5 P91461
92 6 3.1 58 10 Q8RYM0
93 6 3.1 62 10 Q9S8P8
94 6 3.1 67 15 Q71871
95 6 3.1 68 17 Q9S8S2
96 6 3.1 69 2 Q9F812
97 6 3.1 69 10 Q8GXR2
98 6 3.1 71 11 Q8CSN1
99 6 3.1 75 10 Q9LDT2
100 6 3.1 75 17 Q97XET

ALIGNMENTS

RESULT 1
Q8FCP3 PRELIMINARY; PRT; 45 AA.
AC Q8FCP3;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN C4250.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AB016768; AAN82666.1;
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 45 AA; 5384 MW; 989P642744F14BA2 CRC64;
Query Match 4.1%; Score 8; DB 16; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Gaps 0;
Matches 8; Conservative 0; Indels 0;
QY 119 NSEYFL 126
DB 36 NSEYFL 43
RESULT 2
Q8B9W3 PRELIMINARY; PRT; 5020 AA.
AC Q8B9W3;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE RTX toxin, putative.
GN SO4149.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadales; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
Read T.D., Eisen J.A., Seshadri R., Ward N., Mehe B., Clayton R.A.,
Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,

RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.,
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015847; AAN57122.1; -
DR TIGR; SO4149; -
KM Complete proteome.
SQ SEQUENCE 5020 AA; 522577 MW; C90816C507DC01B1 CRC64;

Query Match 4.1%; Score 8; DB 16; Length 5020;
Best Local Similarity 100.0%; Pred. No. 16+02; Mismatches 0; Gaps 0;
Matches 8; Conservative 0; Indels 0;

QY 24 FVLNPDGS 31
DB 1918 FVLNPDGS 1925

RESULT 3
Q8NVLO PRELIMINARY; PRT; 88 AA.
AC Q8NVLO;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Truncated int protein.
GN TRUNCATED INT OR MW1952.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA."
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004829; BAB95817.1; -
DR InterPro: IPR002104; Phage integrase.
DR Pfam: PF00589; Phage_integrase; 1.
KM Complete proteome.
SQ SEQUENCE 88 AA; 10353 MW; 816DCB8C8E9B5070 CRC64;
Query Match 3.6%; Score 7; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 34; Mismatches 0; Gaps 0;
Matches 7; Conservative 0; Indels 0;

QY 33 ILKNIST 39
DB 22 ILKNIST 28

RESULT 4
Q9POS2 PRELIMINARY; PRT; 106 AA.
AC Q9POS2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE HSPC203 (P01019) (Hypothetical protein).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Umbilical cord blood;

RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
 RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.,
 RT "Human full length cDNA cloned from cd34+ stem cells."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary tumor;
 RA Xiao H., Song H., Gao G., Ren S., Chen Z., Han Z.,
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA Strausberg R.;
 RL EMBL: AF151037; AAF6123.1; -;
 DR EMBL: AF226729; AAG09730.1; -;
 DR EMBL: BC001702; AAH01702.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 106 AA; 12293 MW; EB84EE268DF62B07 CRC64;

Query Match 3.6%; Score 7; DB 4; Length 106;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 NKISLES 150
 DB 59 NKISLES 65

RESULT 5
 ID Q8S8D6 PRELIMINARY; PRT; 135 AA.
 AC Q8S8D6;
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Predicted protein.
 GN AT2G20613.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsids.
 NC NCB1_TaxID=3702;
 RX NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Cv. Columbia;
 RA Lin X., Kaul S., Shoa T.P., Fujii C.Y., Shen M., Vannken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Mierman W.C.,
 RA Frazer C.M., Venter J.C.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: ACC007048; NAM15367.1; -;
 DR Pfam: PF04504; DUF573; 1.
 SQ SEQUENCE 135 AA; 14945 MW; 9AEC4B7DC288EB00 CRC64;

Query Match 3.6%; Score 7; DB 10; Length 135;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 184 LNNQESK 190
 DB 56 LNNQESK 62

RESULT 6
 ID P74399 PRELIMINARY; PRT; 135 AA.
 AC P74399;
 DT 01-FEB-1997 (T-EMBLrel. 02, Created)
 DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Hypothetical protein g110268.
 GN SL10268.

OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 NC NCB1_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions."
 RL DNA Res. 3:109-136(1996).
 DR EMBL: D90914; BAA18496.1; -;
 KW Hypothetical protein, complete proteome.
 SQ SEQUENCE 135 AA; 14362 MW; E572DE5AB1ED858C CRC64;

Query Match 3.6%; Score 7; DB 16; Length 135;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 110 DLATTLQ 116
 DB 73 DLATTLQ 79

RESULT 7
 ID P71817 PRELIMINARY; PRT; 139 AA.
 AC P71817;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Steroid isomerase, putative.
 GN RV0760C OR MTCY369.05 OR MT0785.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriinae; Mycobacteriaceae; Mycobacterium.
 NC NCB1_TaxID=1773;
 RX NCB1_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gab S., Barry C.E. II, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z80226; CAB02403.1; -;
 DR EMBL: AE006969; AAK45026.1; -;
 DR TIGR: MT0785; -;
 DR Tuberculist; RV0760C; -;
 DR Interpro; IPR02075; NTF2.
 DR Pfam; PF02136; NTF2; 1.

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 139 AA; 15242 MW; 8291E503C8F07FDD CRC64;

Query Match 3.6%; Score 7; DB 16; Length 139;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 NPDGSGI 33
Db 48 NPDGSGI 54

RESULT 8

Q926T1 PRELIMINARY; PRT; 140 AA.

AC Q926T1; 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein Iln2960.

OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;

RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;

RX PubMed=11679669;

RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Darvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Eutlier K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kunz M., Kunst F., Kurapkat G.,
RA Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstedt G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species."
RL Science 294:849-852 (2001).

DR EMBL; AL596174; CAC98185.1; -
DR ListLib; LIn02960; -
DR InterPro; IPR000835; HTH_Marr.
DR Pfam; PF01047; Marr; 1.
DR PRINTS; PR00598; HTHMARR.

DR SMART; SM00347; HTH_MARR; 1.
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 140 AA; 16142 MW; C913BA5CBDE5D9B8 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 140;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 ILKNIST 39
Db 131 ILKNIST 137

RESULT 9

Q8Y3K8 PRELIMINARY; PRT; 140 AA.

AC Q8Y3K8; 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein Iln02827.

GN LMO2827.

OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OX NCBI_TaxID=1639;

RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=EGD-e / Serovar 1/2a;
RA MEDLINE=21537279; PubMed=11679669;

RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Darvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Eutlier K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kunz M., Kunst F., Kurapkat G.,
RA Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstedt G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species."
RL Science 294:849-852 (2001).

DR EMBL; AL591984; CAD01040.1; -
DR ListLib; LMO02827; -
DR InterPro; IPR000835; HTH_Marr.
DR Pfam; PF01047; Marr; 1.
DR PRINTS; PR00598; HTHMARR.

DR SMART; SM00347; HTH_MARR; 1.
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 140 AA; 16177 MW; EB08BAFDD0014F34 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 140;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 ILKNIST 39
Db 131 ILKNIST 137

RESULT 10

Q9BGP3 PRELIMINARY; PRT; 142 AA.

AC Q9BGP3; 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 15.6 kDa protein.

OS Macaca fascicularis (Crib eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.

OX NCBI_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Frontal cortex;

RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,

RA Suzuki Y., Sugano S., Hashimoto K.,
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries."

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB056429; BAB33087.1; -

DR HSSP; P15374; IUCH.

DR MEROPS; C12.001; -

DR InterPro; IPR001578; UCH_1.

DR Pfam; PF01088; Peptidase_C12; 1.

DR PRINTS; PR00707; UBCTHYDLASE.

DR PROSITE; PS00140; UCH_1; 1.

KW Hypothetical protein.
SQ SEQUENCE 142 AA; 15632 MW; 2D9F63762DA54D71 CRC64;

Query Match 3.6%; Score 7; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 CFEKNEA 90
Db 51 CFEKNEA 57

```

RESULT 11
ID 018355 PRELIMINARY; PRT; 153 AA.
AC 018355;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Short form of CHIP.
GN CHIP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97477378; PubMed=9334334;
RA Morcillo P., Rosen C., Baylies M.K., Dorsett D.;
RT "Chip, a widely expressed chromosomal protein required for
RT segmentation and activity of a remote wing margin enhancer in
RT Drosophila.";
RL Genes Dev. 11:2729-2740(1997).
DR EMBL: AF010327; AAB62576.1; -.
SQ SEQUENCE 153 AA; 15028 MW; F6006B2289A80B52 CRC64;

Query Match 3.6%; Score 7; DB 5; Length 153;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 VLNPDGS 31
DB 40 VLNPDGS 46

RESULT 12
Q9WS22 PRELIMINARY; PRT; 164 AA.
ID Q9WS22;
AC Q9WS22;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Glycoprotein gp82/105 (Fragment).
GN U98.
OS Human herpesvirus 6.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxId=10368;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HST;
RA Iseegawa Y.;
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=HST;
RA MEDLINE=99412319; PubMed=10482554;
RA Iseegawa Y., Mukai T., Nakano K., Kagawa M., Chen J., Mori Y.,
RA Yamashita K., Kawanishi K., Sashihara J., Hata A., Zou P., Kosuge H.,
RT "Comparison of the complete DNA sequences of human herpesvirus 6
RT variants A and B.";
RL J. Virol. 73:8053-8063(1999).
DR EMBL: AB021506; BAA78315.1; -.
FT NON_TER 1
SQ SEQUENCE 164 AA; 18833 MW; E8E23BDE627B6BCD CRC64;

Query Match 3.6%; Score 7; DB 12; Length 164;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 TLQSSNS 120
DB 126 TLQSSNS 132

```

```

RESULT 13
ID 08V980 PRELIMINARY; PRT; 177 AA.
AC 08V980;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE FOF1-ATPase subunit delta (ATP synthase delta chain).
GN ATPH.
OS Colwellia maris (Vibrio sp. (strain ABE-1)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Colwellia.
OX NCBI_TaxId=77524;
RN [1]
RP SEQUENCE FROM N.A.
RX Takada Y., Takiya S.;
RA "Colwellia maris atp operon, complete sequence.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RL FUNCTION: THIS PROTEIN SEEMS TO BE PART OF THE STALK THAT LINKS
CC CF(0) TO CF(1). IT EITHER TRANSMITS CONFORMATIONAL CHANGES FROM
CC CF(0) INTO CF(1) OR IS IMPLICATED IN PROTON CONDUCTION (BY
CC SIMILARITY).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATPASE DELTA CHAIN FAMILY.
DR EMBL: AB035129; BAB82481.1; -.
DR InterPro: IPR000711; ATPsyn_OSCP.
DR Pfam: PF0213; OSCP. 1.
DR PRINTS: PR00125; ATPASEDELTA.
DR TIGRPFAM: TIGR01145; ATP_synth_delta; 1.
DR PROSITE: PS00389; ATPASE_DELTA; 1.
DR ATP synthase; CF(1); Hydrolyase.
SQ SEQUENCE 177 AA; 19131 MW; 1CF62614457A0A78 CRC64;

Query Match 3.6%; Score 7; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 LATLQGS 117
DB 171 LATLQGS 177

RESULT 14
ID 018354 PRELIMINARY; PRT; 182 AA.
AC 018354;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE Short form of CHIP.
GN CHIP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97477378; PubMed=9334334;
RA Morcillo P., Rosen C., Baylies M.K., Dorsett D.;
RT "Chip, a widely expressed chromosomal protein required for
RT segmentation and activity of a remote wing margin enhancer in
RT Drosophila.";
RL Genes Dev. 11:2729-2740(1997).
DR EMBL: AF010326; AAB62575.1; -.
SQ SEQUENCE 182 AA; 19474 MW; B4C12B1B167A78AE CRC64;

Query Match 3.6%; Score 7; DB 5; Length 182;

```

Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 VLNPDGS 31
Db 40 VLNPDGS 46

RESULT 15

08X563 ID 08X563 PRELIMINARY; PRT; 187 AA.
AC 08X563;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative major fibrial subunit (Putative fibrial-like protein).
GN Z3601 OR EC3222.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grodbeck E.V., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
DR EMBL: AE005465; AAGS7467.1; -
DR EMBL: AF002561; BAB36645.1; -
DR InterPro: IPR000259; FimDrial.
DR Pfam: PF00419; FimDrial.
KM Complete proteome.
SQ SEQUENCE 187 AA; 19661 MW; 03BC62C650288134 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 187;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 GSGILKN 36
Db 60 GSGILKN 66

RESULT 16

097KC7 ID 097KC7 PRELIMINARY; PRT; 203 AA.
AC 097KC7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Hypothetical protein CAC0992.
GN CAC0992
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_Taxid=1488;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hiltz J., Wolf Y.I.,
RA Tatusev R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett-G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum";
RL J. Bacteriol. 183:4823-4838(2001).
RW EMBL: AE007615; AAK78968.1; -
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 203 AA; 23809 MW; 41CB5CB3BFD0D3FC CRC64;

Query Match 3.6%; Score 7; DB 16; Length 203;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 FEKNEAH 91
Db 63 FEKNEAH 69

RESULT 17

08E919 ID 08E919 PRELIMINARY; PRT; 210 AA.
AC 08E919;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Exopolysaccharide synthesis protein, putative.
GN SO4288.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_Taxid=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297666; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seehardt R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsaplin A., Scott J., Beaman M., Binkac L., Daugherty S.,
RA DeGory R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL: AE015861; AAN57257.1; -
DR TIGR: SO4288; -
KM Complete proteome.
SQ SEQUENCE 210 AA; 23150 MW; F75DFC3A0F9847B4 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 210;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 FLUPISI 12
Db 53 FLUPISI 59

RESULT 18

08ZTK9 ID 08ZTK9 PRELIMINARY; PRT; 218 AA.
AC 08ZTK9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

```

DE Hypoetical protein PAE3205.
GN PAE3205.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL: AE009917; AAL64752.1; -.
KW Hypoetical protein; Complete proteome.
SQ SEQUENCE 218 AA; 23395 MW; 10C60FC98973C44 CRC64;

Query Match 3.6%; Score 7; DB 17; Length 218;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRPISIS 13
DB 47 LRPISIS 53

RESULT 19
Q8E2P2 PRELIMINARY; PRT; 223 AA.
AC Q8E2P2;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypoetical protein.
GN GBS2080.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=2242508; PubMed=12354221;
RA Glaeser P., Rusniok C., Buchrieser C., Chevalier F., Frangoul L.,
RA Weader T., Zouine M., Couve E., Lailou L., Poyart C., Titeu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1489-1513(2002).
DR EMBL: AL766856; CAD47739.1; -.
KW Sagalict; gbs2080; Complete proteome.
SQ SEQUENCE 223 AA; 24992 MW; E0566F4903564CE1 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 223;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 LKNISTK 40
DB 72 LKNISTK 78

RESULT 20
Q8DMU2 PRELIMINARY; PRT; 223 AA.
AC Q8DMU2;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypoetical protein.

```

```

GN SAG2121.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Messela M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beaman M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouli H., Mulligan S.,
RA Carey H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Jacobini E.T., Breltoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL: AE014287; AAN00979.1; -.
DR TIGR: SAG2121; -.
KW Hypoetical protein; Complete proteome.
SQ SEQUENCE 223 AA; 24992 MW; E0566F4903564CE1 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 223;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 LKNISTK 40
DB 72 LKNISTK 78

RESULT 21
Q9JYW4 PRELIMINARY; PRT; 242 AA.
AC Q9JYW4;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Hypoetical protein NMB1404.
GN NMB1404.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=2015755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gilm M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citron H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouli H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL: AE002489; AAF41767.1; -.
DR TIGR: NMB1404; -.
KW Hypoetical protein; Complete proteome.
SQ SEQUENCE 242 AA; 28657 MW; C5533BAF53FF381C CRC64;

Query Match 3.6%; Score 7; DB 16; Length 242;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 FLIPISI 12

```

Db 64 FLIPISI 70

RESULT 22

Q9JTV0 PRELIMINARY; PRT; 242 AA.
AC Q9JTV0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative integral membrane protein.
GN NMA1624.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C., Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T., Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S., Jagers K., Leather S., Moule S., Mungall K., Quail M.A., Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J., Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491."
RL Nature 404:502-506(2000).
DR EMBL; AL162756; CAB84852.1; -
KW Complete proteome.
SQ SEQUENCE 242 AA; 28701 MW; FE2326C8DD2DED3 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 242;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 FLIPISI 12
Db 64 FLIPISI 70

RESULT 23

O8E0N9 PRELIMINARY; PRT; 257 AA.
AC O8E0N9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Membrane protein, putative.
GN SAG0689.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=1200547;
RA Tettelin H., Mesigian V., Cieslewicz M.J., Eisen J.A., Peterson S., Messiss M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D., Madoff L.C., Wolf A.M., Beaman M.J., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R., Radune D., Fedorova N.B., Scanlan D., Khoult H., Mulligan S., Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M., Jacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D., Rinaldo D., Rappunli R., Telford J.L., Kasper D.L., Grandi G., Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae."
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
EMBL; AE014222; AAM99577.1; -

DR TIGR; SAG0689; -
KW Complete proteome.
SQ SEQUENCE 257 AA; 28036 MW; B6B8ABF0213DC5B CRC64;

Query Match 3.6%; Score 7; DB 16; Length 257;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 SGENSIS 157
Db 2 SGENSIS 8

RESULT 24

O9CKK9 PRELIMINARY; PRT; 275 AA.
AC O9CKK9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein PM1603.
GN PM1603.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=2145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006198; AAK03687.1; -
KW Hypothetical protein, Complete proteome.
SQ SEQUENCE 275 AA; 32005 MW; EC507D6DA76FD36 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 275;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 LKNISTK 40
Db 147 LKNISTK 153

RESULT 25

P88930 PRELIMINARY; PRT; 278 AA.
AC P88930;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ORF 42.
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=37296;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97121480; PubMed=8962146;
RA Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Madgalena D., Parry J.P., Penzuzi D., Edelman I.S., Chang Y., Moore P.S.;
RT "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8)."
RL Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97094384; PubMed=8939871;
RA Moore P.S., Bashoff C., Weiss R.A., Chang Y.;
RT "Molecular mimicry of human cytokine and cytokine response pathway genes by KSHV."
RL Science 274:1739-1744(1996).

```

RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97296220; PubMed=9151804;
RA Neipel F., Albrecht J.C., Fleckenstein B.;
RT "Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus
RT human herpesvirus 8: determinants of its pathogenicity?";
RL J. Virol. 71:4187-4192 (1997).
DR EMBL: U75698; AAC57124.1; -.
DR InterPro: IPR002600; Herpes_UL7.
DR Pfam: PF01677; Herpes_UL7.1.
SQ SEQUENCE 278 AA; 31225 MW; 1206A1893811EE46 CRC64;

Qy Query Match 3.6%; Score 7; DB 12; Length 278;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 148 LESSGEN 154
223 LESSGEN 229

RESULT 26
ID 006058 PRELIMINARY; PRT; 285 AA.
AC 006058;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Chromosome XII COSMID 8084.
GN YLR404W OR L8084.3.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansgore W.,
RA Bente V., Bruckner M., Deline H., Dubois E., Duersthoft A.,
RA Entian K.D., Floeth M., Goffeau A., Hedling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hliger F., Kleine K., Kotter P.,
RA Louis E.J., Messenguy F., Mewes H.W., Miosga T., Mostl D.,
RA Muller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purrelle B., Rechemann S., Rieger M., Rinke M., Rose M.,
RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrutiarazu L.A., Vandenbol M., Verhasselt P.,
RA Vlerendeels F., Voet M., Volckaert G., Voss H., Wambuit R., Medler E.,
RA Medler H., Zimmermann F.K., Zollner A., Hant J., Hohnheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:0-0(0).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RA Du Z.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RA Waterston R.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RA Cherry J.M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U19729; AAB8342.1; -.
DR SGD: S0004396; YLR404W.
SQ SEQUENCE 285 AA; 32590 MW; 3966CB0967B7B62B CRC64;

Query Match 3.6%; Score 7; DB 3; Length 285;
Best Local Similarity 100.0%; Pred. No. 94;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 KISLESS 151
Db 195 KISLESS 201

RESULT 27
ID 08NMK8 PRELIMINARY; PRT; 294 AA.
AC 08NMK8;
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Succinyl-CoA synthetase alpha subunit (EC 6.2.1.5).
GN CGL265.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF005282; BAB99958.1; -.
DR InterPro: IPR003781; CoA binding.
DR InterPro: IPR005811; CoA_ligase.
DR InterPro: IPR005810; CoA_lig_alpha.
DR Pfam: PF02629; CoA_binding_1.
DR Pfam: PF00549; ligase-CoA; 1.
DR TIGRFAMs: TIGR01019; succoalalpha; 1.
DR PROSITE: PS01216; SUCCINYL_COA_LIG_1; 1.
DR PROSITE: PS00399; SUCCINYL_COA_LIG_2; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 294 AA; 30259 MW; BD9198F23E3C80AA CRC64;

Query Match 3.6%; Score 7; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 KHALEAV 71
Db 266 KHALEAV 272

RESULT 28
ID 09V263 PRELIMINARY; PRT; 303 AA.
AC 09V263;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE Phosphate regulatory protein, putative.
GN PAB0144.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ248283; CAB49135.1; -.
DR InterPro: IPR002754; Phou.
DR Pfam: PF01895; Phou; 1.
DR ProDom: PD014067; Phou; 1.
KW Complete proteome.
SQ SEQUENCE 303 AA; 35250 MW; 643EE120BDA1466D CRC64;

```


Query Match 3.6%; Score 7; DB 17; Length 303;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 VLNPDGS 31
Db 37 VLNPDGS 43

RESULT 29

Q9VNB4 PRELIMINARY; PRT; 308 AA.
AC Q9VNB4
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CG2663 protein.
GN CG2663.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazet R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-U., Andrews-Pfannkuch C., Baldwin D.,
RA Balise R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Buser D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Esler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matvei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster.",
RL Science 287:2185-2195(2000).
DR EMBL; AEO03603; AAF52032.1; -
DR FLYBase; FBgn0037323; CG2663.
DR InterPro; IPR001251; CRAL-TRIO.
DR InterPro; IPR001071; Retbind/locotrans.
DR Pfam; PF00650; CRAL_TRIO; 1.
DR PRINTS; PR00180; CRETINALDHP.
DR SMART; SMO0516; SEC14; 1.

DR PROSITE; PSS0191; CRAL-TRIO; 1.
SQ SEQUENCE 308 AA; 35512 MW; FBE7A3AC43D472EE CRC64;

Query Match 3.6%; Score 7; DB 5; Length 308;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GDIFILD 63
Db 156 GDIFILD 162

RESULT 30

Q9NUP8 PRELIMINARY; PRT; 315 AA.
AC Q9NUP8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ11218.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Nagatsuma T., Hosokori T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakami K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagai K., Masuno Y.,
RA Nishimura K., Iwayanagi T.;
RT "NEBO human cDNA sequencing project";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK02080; BAA92073.1; -
RX Hypothetical protein.
SQ SEQUENCE 315 AA; 34949 MW; 8C1EBF93073E4690 CRC64;

Query Match 3.6%; Score 7; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 LEAVSYS 74
Db 185 LEAVSYS 191

RESULT 31

Q81099 PRELIMINARY; PRT; 315 AA.
AC Q81099
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG2663-PA (RE58223p).
GN CG2663.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazet R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.-U., Andrews-Pfannkuch C., Baldwin D.,

RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Borkhan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshell A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclet J.M.,
 RA Palazzolo M., Pictman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskae R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodgett M., Worley K.C., Wu D., Yang S., Yao O.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L., Ye J.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Goeys J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dreese D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Matzel B., Moshell A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Paclet J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pictman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskae R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of *Drosophila melanogaster* genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Croebly M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminler J.S., Prochuk S.B., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Seare S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
 RA Aeburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
 RT "Annotation of *Drosophila melanogaster* genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA STRAIN=y;
 RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Dreese D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunoo J., Paclet J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.,
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO03603; AAN13281.1; -
 DR EMBL: BT001686; AAN71441.1; -
 SQ SEQUENCE 315 AA; 36210 MW; C30D1DBFF394B3FD CRC64;
 Query Match 3.6%; Score 7; DB 5; Length 315;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 57 GDIPLD 63
 Db 163 GDIPLD 169
 RESULT 32
 ID Q81SS3 PRELIMINARY; PRT; 335 AA.
 AC Q81SS3;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DE Acyl-CoA Z/E11 desaturase.
 OS Chorisoneura rosaceana (Oblique banded leafroller).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Tortricidae; Tortricidae; Tortricidae; Tortricinae; Chorisoneura.
 OX NCBI_TaxID=27543;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hao G., O'Connor M., Liu W., Roelofs W.L.,
 RT "Cloning and expressing of desaturase genes from oblique banded
 RT leafroller moths, *Chorisoneura rosaceana*."
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF545481; AAN41250.1; -
 SQ SEQUENCE 335 AA; 38784 MW; 7DAB84EF770183F6 CRC64;
 Query Match 3.6%; Score 7; DB 5; Length 335;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 99 AENKMTF 105
 Db 246 AENKMTF 252
 RESULT 33
 ID Q9EWX3 PRELIMINARY; PRT; 340 AA.
 AC Q9EWX3;
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
 DE 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Putative integral membrane protein.
 GN SC01753 OR 25C134.06.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=A3(2);
 RC Saunders D.C., Harris D.,
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=A3(2);
 RC Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.,
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapapete D., Eichner A., Cullum J.,
 RA Knaeth H., Hopwood D.A., and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RL coelicolor A3(2)." / (2002).
 RL Nature 417:141-147(2002).
 DR EMBL: AL939110; CAC12926.1; -;
 DR InterPro: IPR00326; PA_PTPase.
 DR Pfam: PF01569; PAP2; 1.
 DR SMART: SM00014; acidPpc; 1.
 KW Complete proteome.
 SQ SEQUENCE 340 AA; 37780 MW; 347642B1DF9A9B2 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 340;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 WLGGDIF 60
 Db 154 WLGGDIF 160

RESULT 34
 O9BS33 PRELIMINARY; PRT; 342 AA.
 AC O9BS33;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Similar to hypothetical protein FLJ11218.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC005364; AA05364.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 342 AA; 37833 MW; C598E09414E7D893 CRC64;

Query Match 3.6%; Score 7; DB 4; Length 342;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 LEAVSYS 74
 Db 185 LEAVSYS 191
 RESULT 35
 O961T4 PRELIMINARY; PRT; 357 AA.
 AC O961T4;

DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE GH089489.
 GN MT11 OR CG10569 OR CG32417.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Broketein P., Hong L., Agbayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunco J., Pieble J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY047577; AAK7309.1; -;
 DR FLYBase: FBgn0040298; Myt1.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 357 AA; 40706 MW; 4ACB07B97F245176 CRC64;

Query Match 3.6%; Score 7; DB 5; Length 357;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 143 INKISLE 149
 Db 141 INKISLE 147

RESULT 36
 O18356 PRELIMINARY; PRT; 365 AA.
 AC O18356;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Short form of CHIP.
 GN CHIP.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97477378; PubMed=9334334;
 RA Morcillo P., Rosen C., Bayliss M.K., Dorsett D.;
 RT "chip, a widely expressed chromosomal protein required for
 RT segmentation and activity of a remote wing margin enhancer in
 RT Drosophila.";
 RL Genes Dev. 11:2729-2740(1997).
 DR EMBL: AF010328; AAB62577.1; -;
 DR InterPro: IPR002691; LIM_bind.
 DR Pfam: PF01803; LIM_bind; 1.
 SQ SEQUENCE 365 AA; 39437 MW; F83187B002EC1566 CRC64;

Query Match 3.6%; Score 7; DB 5; Length 365;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 VLNPDGS 31
 Db 40 VLNPDGS 46

```

RESULT 37
ID Q8N998 PRELIMINARY; PRT: 374 AA.
AC Q8N998;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ36159.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Suzuki O., Sasaki N., Aotaka S., Shoji T., Ichihara T., Shiohara N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsumura H., Takahashi E., Katsuka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Matanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Seto K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Ohima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuhio Y., Nagai K., Isogai T.,
RT "NEO human cDNA sequencing project."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC095478; BAC04556.1; -.
KW Hypothetical protein.
SQ SEQUENCE 374 AA; 43808 MW; 161160AC6AC1491 CRC64;

Query Match 3.6%; Score 7; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 DRKHALE 69
DB 306 DRKHALE 312

RESULT 38
ID Q8PCW8 PRELIMINARY; PRT: 376 AA.
AC Q8PCW8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GTP cyclohydrolase II/3, 4-dihydroxy-2-butanone 4-phosphate
DE synthase.
GN RIBA OR XCC0696.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OC NCB1_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 33913 / NCPPB 528;
RA MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria U.B., Ferreira A.J.S., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,

```

```

RA Trindade dos Santos M., Truffi D., Teal S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463 (2002).
DR EMBL; AE012168; AAM40012.1; -.
DR InterPro; IPR000422; DHP_synthase.
DR InterPro; IPR000926; GTP_cyclohydrolII.
DR Pfam; PF00926; DHP_synthase; 1.
DR Pfam; PF00925; GTP_cyclohydrol2; 1.
DR ProDom; PD003034; DHP_synthase; 1.
DR TIGRFAMs; TIGR00506; r1bB; 1.
DR TIGRFAMs; TIGR00506; r1bB; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 376 AA; 40803 MW; F9B5A7DA4B706798 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 VLNPDGS 31
DB 165 VLNPDGS 171

RESULT 39
ID Q8PPD7 PRELIMINARY; PRT: 379 AA.
AC Q8PPD7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GTP cyclohydrolase II.
GN RIBA OR XAC0749.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OC NCB1_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=306 / ATCC 13902 / XV 101;
RA MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria U.B., Ferreira A.J.S., Ferreira R.C.C., Ferro W.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madela A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teal S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463 (2002).
DR EMBL; AE011704; AAM35638.1; -.
DR InterPro; IPR000422; DHP_synthase.
DR InterPro; IPR000926; GTP_cyclohydrolII.
DR Pfam; PF00925; DHP_synthase; 1.
DR ProDom; PD003034; DHP_synthase; 1.
DR TIGRFAMs; TIGR00506; r1bB; 1.
DR TIGRFAMs; TIGR00506; r1bB; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 379 AA; 41176 MW; E88FF2AD596442E CRC64;

Query Match 3.6%; Score 7; DB 16; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 25 VLNPDS 31
 Db 165 VLNPDS 171

RESULT 40

Q9XDF8 PRELIMINARY; PRT; 425 AA.
 AC Q9XDF8;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE NonK.
 GN NONK.
 OS Streptomyces griseus subsp. griseus.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 CX NCBI_TaxID=67263;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ETH A7796;
 RA MEDLINE=20117817; PubMed=1065022;
 RA Walczak R.J., Woo A.J., Ström W.R., Priestley N.D.,
 RT "Nonactin biosynthesis: the potential nonactin biosynthesis gene
 cluster contains type II polyketide synthase-like genes."
 RL FEMS Microbiol. Lett. 183:171-175 (2000).
 DE EMBL: AF074603; AD37450.1; -;
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR Pfam: PF00109; Ketoacyl-synt. 1.
 DR Pfam: PF02801; Ketoacyl-synt. C; 1.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.
 SQ SEQUENCE 425 AA; 44410 MW; D0176DF9D265926 CRC64;

Query Match 3.6%; Score 7; DB 2; Length 425;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 DAKSLFV 25
 Db 123 DAKSLFV 129

RESULT 41

O8N007 PRELIMINARY; PRT; 429 AA.
 AC O8N007;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Histidyl-tRNA synthetase (EC 6.1.1.21).
 GN CGL1647.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 CX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
 RL EMBL: AP005279; BAB99040.1; -;
 DR InterPro: IPR004154; HGP anticonodon.
 DR InterPro: IPR004516; H18S.
 DR InterPro: IPR002314; tRNA-synt_2b.
 DR InterPro: IPR006195; tRNA_Ligase_II.
 DR Pfam: PF03129; HGP anticonodon; 1.
 DR Pfam: PF00587; tRNA-synt_2b; 1.
 DR TIGRfams: TIGR00442; h18S; 1.
 DR PROSITE: PS50862; AA_tRNA_Ligase_II; 1.
 KW Aminoacyl-tRNA synthetase; Complete proteome.
 SQ SEQUENCE 429 AA; 47126 MW; 06075626963CDEB CRC64;

Query Match 3.6%; Score 7; DB 16; Length 429;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 YTKTCFE 86
 Db 271 YTKTCFE 277

RESULT 42

O8D6L6 PRELIMINARY; PRT; 446 AA.
 AC O8D6L6;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Sugar phosphatase permealase.
 GN VV20513.
 OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 CX NCBI_TaxID=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMCP6;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.E.;
 RT "Complete genome sequence of Vibrio vulnificus CMCP6."
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DE EMBL: AE016809; AAC07463.1; -;
 KW Complete proteome.
 SQ SEQUENCE 446 AA; 49487 MW; C085D0C7DB4987B7 CRC64;

Query Match 3.6%; Score 7; DB 16; Length 446;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 GFSSSLI 173
 Db 113 GFSSSLI 119

RESULT 43

O9DA19 PRELIMINARY; PRT; 450 AA.
 AC O9DA19;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE 1700023B02Rik protein.
 GN 1700023B02Rik.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilmig L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohno S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nucleic Acids Res. 29:685-690(2001).
 DR EMBL; AK006260; BAB24488.1; -
 DR MGD; AK0194185; 1700023B02Rik.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR PROSITE; PS00190; CYTOCHROME_C_1.
 SQ SEQUENCE 450 AA; 51852 MW; F32F11BE6D6A4EAC CRC64;

Query Match 3.6%; Score 7; DB 11; Length 450;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 ESSGENS 155
 DB 356 ESSGENS 362

RESULT 44
 OSJRN2 PRELIMINARY; PRT; 472 AA.
 ID Q9JRN2;
 AC Q9JRN2;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE Putative glycosyltransferase.
 OS Actinobacillus actinomycetemcomitans (Haemophilus
 OS actinomycetemcomitans).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 NC NCB1_TaxID=714;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SUNYAB 75;
 RA Suzuki N., Nakano Y., Yoshida Y., Nako H., Yamashita Y., Koga T.;
 RT "Genetic analysis of the gene cluster for the synthesis of serotype a-
 RT specific polysaccharide antigen in A. actinomycetemcomitans";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB046360; BAB03211.1; -
 DR InterPro: IPR003362; Bact_transf.
 DR Pfam; PF02397; Bact_transf; 1.
 KW Transferrase.
 SQ SEQUENCE 472 AA; 55262 MW; C2FESAFD16CB9FD0 CRC64;

Query Match 3.6%; Score 7; DB 2; Length 472;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 KINKISL 148
 DB 8 KINKISL 14

RESULT 45
 OSJN76 PRELIMINARY; PRT; 477 AA.
 ID Q9JN76;
 AC Q9JN76;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
 DE Hypothetical 51.0 kDa protein.
 GN F1P2.170.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucotyledon II; Brassicales; Brassicaceae; Arabidopsids.
 NC NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Choline N., Robert C., Brothier P., Wincker P., Catolico L.,

RA Artiguenave F., Saurin W., Weissenbach J., Newes H.W., Lemcke K.,
 RA Mayer K.F.X., Queller F., Salanoubat M.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL132955; CAB61988.1; -
 DR InterPro: IPR005333; TCP.
 DR Pfam; PF03634; TCP; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 477 AA; 50987 MW; C0049F717D7B395 CRC64;

Query Match 3.6%; Score 7; DB 10; Length 477;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 TLQSSNS 120
 DB 296 TLQSSNS 302

RESULT 46
 OSJUR9 PRELIMINARY; PRT; 488 AA.
 ID Q9JUR9;
 AC Q9JUR9;
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Transcription factor hunchback (Fragment).
 GN HUNCHBACK.
 OS Schistocerca americana (American grasshopper).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;
 OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
 NC NCB1_TaxID=7009;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21450564; PubMed=11566852;
 RA Patel N.H., Hayward D.C., Tall S., Pirk N.R., DiPietro D., Ball E.E.;
 RT "Grasshopper hunchback expression reveals conserved and novel aspects
 RT of axis formation and segmentation."
 RL Development 128:3459-3472(2001).
 DR EMBL; AY040605; AAK84959.1; -
 DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam; PF00096; Zf-C2H2; 4.
 DR SMART; SM00355; Znf_C2H2; 6.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 4.
 KW Metal-binding; Zinc; Zinc-finger.
 FT NON_TER 1.
 SQ SEQUENCE 488 AA; 53734 MW; FFD5D477A5128376 CRC64;

Query Match 3.6%; Score 7; DB 5; Length 488;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 VLNPDGS 31
 DB 120 VLNPDGS 126

RESULT 47
 OSJ3Z00 PRELIMINARY; PRT; 489 AA.
 ID Q9J3Z00;
 AC Q9J3Z00;
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
 DE AT3G47620/F1P2.170.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosida II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shinn P., Banh J.,
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis ORF clones."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY058874; AL124261.1; -
 DR EMBL: AY103304; AA65356.1; -
 DR InterPro: IPR005333; TCP.
 DR Pfam: PF03634; TCP; 1.
 SQ SEQUENCE 489 AA; 52300 MW; AC777E2D96FC7586 CRC64;

Query Match 3.6%; Score 7; DB 10; Length 489;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 114 TLQSSNS 120
 Db 308 TLQSSNS 314

RESULT 48
 Q95US0 PRELIMINARY; PRT; 512 AA.
 AC Q95US0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hunchback transcription factor (Fragment).
 GN HB.
 OS Locusta migratoria (Migratory locust).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthoptera; Orthoptera; Caelifera; Acridoidea;
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.
 OX NCBI_TaxID=7004;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21450564; PubMed=11566852;
 RA Patel N.H., Hayward D.C., Lall S., Pirkl N.R., DiPietro D., Ball E.E.;
 RT "Grashopper hunchback expression reveals conserved and novel aspects
 of axis formation and segmentation."
 RL Development 128:3459-3472(2001).
 DR EMBL: AY040538; AK81873.1; -
 DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam: PF00096; Zf-C2H2; 5.
 DR SMART: SM00355; Znf_C2H2; 6.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 4.
 KW Metal-binding; zinc; zinc-finger.
 FT NON_TER 1
 SQ SEQUENCE 512 AA; 55962 MW; 0E0862E6B4B4A15 CRC64;
 Query Match 3.6%; Score 7; DB 5; Length 512;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 25 VLNPDS 31
 Db 120 VLNPDS 126

RESULT 49
 Q8VD45 PRELIMINARY; PRT; 531 AA.
 AC Q8VD45;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE UDP-glucuronosyltransferase 1A5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RA Kessler F.K., Kessler M.R., Auyeung D.J., Ritter J.K.;
 RT "Glucuronidation of acetaminophen catalyzed by multiple rat phenol
 UDP-glucuronosyltransferases."
 RL Drug Metab. Dispos. 0:0-0(2002).
 DR EMBL: AF461734; AL67850.1; -
 DR InterPro: IPR002213; UDP_gluc_trans.
 DR Pfam: PF00201; UDPGT; 1.
 DR PROSITE: PS00375; UDPGT; 1.
 KW Transferase.
 SQ SEQUENCE 531 AA; 60013 MW; A18B5C6AE5B07FA CRC64;

Query Match 3.6%; Score 7; DB 11; Length 531;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 33 ILKNIIST 39
 Db 113 ILKNIIST 119

RESULT 50
 Q9VRP6 PRELIMINARY; PRT; 533 AA.
 AC Q9VRP6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG32417 protein.
 GN MYT1 OR CG10569 OR CG32417.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benoe P.V., Berman B.P., Bhandari D., Bolhakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Houston D., Houston K.A., Howard T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mleishina N.V., Mobarry C., Morris J., Mostrel A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murzy D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon S., Nusser D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ceiniker S.E., Adams M.D., Krommler B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Benson J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferrera S., Frise E., Galie R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Ceiniker S.E.,
 RA Ciamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommler B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FLYBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN EMBL: ABR03565; AARS0747.3; -
 DR FLYBase; FBgn0040298; Myt1.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD0000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 533 AA; 61131 MW; 3DCD3458914D0645 CRC64;
 Query Match 3.6%; Score 7; DB 5; Length 533;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 143 INKISLE 149
 DB 317 INKISLE 323

Search completed: January 28, 2004, 18:22:38
 Job time : 82 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: January 28, 2004, 15:56:06 ; Search time 399 Seconds
(without alignments)
8950.768 Million cell updates/sec

Title: US-09-881-556A-3

Perfect score: 1323
Sequence: 1 gaattcgagcagagatttt.....aaaaaaaaaacctcgag 1323

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 15

Total number of hits satisfying chosen parameters: 144779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

1: N_Geneseq_19Jun03.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
26: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1323	100.0	1323	24	ABA91278	Cryptosporidium pa
2	29	2.2	549	19	AAV38762	Complementing stru
3	29	2.2	1049	19	AAV38761	CDNA nfe2p18-549 en
4	29	2.2	1049	19	AAV38761	Human secreted pro
5	29	2.2	1049	24	AB573545	Human cDNA #1 for
6	29	2.2	1108	24	AB573545	Human polynucleoti
7	29	2.2	1198	20	AAK2260	Human secreted pro
8	28	2.1	743	21	AAK21722	Human breast and o

9	28	2.1	2150	21	AA15648	Human prostate can
10	27	2.0	300	19	AAV42737	CDNA sequence of I
11	27	2.0	559	22	ABA14083	Human nervous syst
12	27	2.0	708	22	AA535580	Human cardiovascular
13	27	2.0	740	21	AA001696	Human colon cancer
14	27	2.0	1017	18	AA763042	Aspergillus niger
15	27	2.0	1017	18	AA760426	Cellulase cDNA c10
16	27	2.0	1061	20	AA207191	Human lung tumour
17	27	2.0	1061	21	AA279143	Human lung tumour-
18	27	2.0	1061	23	AA232128	Human lung tumour-
19	27	2.0	1249	20	AA215856	Human gene express
20	27	2.0	1414	23	AA565294	DNA encoding novel
21	27	2.0	1427	24	ABU90278	Human polynucleoti
22	27	2.0	1553	21	AA279131	Human cancer assoc
23	27	2.0	2509	21	AA279131	Human secreted pro
24	27	2.0	2528	21	AA279131	Human secreted pro
25	27	2.0	2785	24	ABU01572	Human secreted pro
26	26	2.0	309	20	AA222554	Human secreted pro
27	26	2.0	317	20	AA207205	Human lung tumour
28	26	2.0	317	21	AA279062	Human lung tumour-
29	26	2.0	317	23	AA231337	Human lung tumour-
30	26	2.0	352	21	AA231337	Human secreted pro
31	26	2.0	385	21	AA295544	Human secreted pro
32	26	2.0	403	18	AA784528	Human prostate pr
33	26	2.0	403	20	AA235858	Human secreted pro
34	26	2.0	404	21	AA235858	Human secreted pro
35	26	2.0	494	21	AA235858	Human secreted pro
36	26	2.0	519	22	AA235858	Human secreted pro
37	26	2.0	519	22	AA235858	Human secreted pro
38	26	2.0	524	22	AA235858	Human secreted pro
39	26	2.0	559	21	AA235858	Human secreted pro
40	26	2.0	577	22	AA235858	Human secreted pro
41	26	2.0	591	22	AA235858	Human secreted pro
42	26	2.0	639	22	AA235858	Human secreted pro
43	26	2.0	646	22	AA235858	Human secreted pro
44	26	2.0	655	25	AA235858	Human secreted pro
45	26	2.0	655	25	AA235858	Human secreted pro
46	26	2.0	682	22	AA235858	Human secreted pro
47	26	2.0	695	21	AA235858	Human secreted pro
48	26	2.0	712	20	AA235858	Human secreted pro
49	26	2.0	735	21	AA235858	Human secreted pro
50	26	2.0	744	21	AA235858	Human secreted pro
51	26	2.0	786	19	AA235858	Human secreted pro
52	26	2.0	815	25	AA235858	Human secreted pro
53	26	2.0	829	25	AA235858	Human secreted pro
54	26	2.0	829	25	AA235858	Human secreted pro
55	26	2.0	929	21	AA235858	Human secreted pro
56	26	2.0	929	21	AA235858	Human secreted pro
57	26	2.0	1001	24	AA235858	Human secreted pro
58	26	2.0	1046	22	AA235858	Human secreted pro
59	26	2.0	1046	25	AA235858	Human secreted pro
60	26	2.0	1046	25	AA235858	Human secreted pro
61	26	2.0	1051	21	AA235858	Human secreted pro
62	26	2.0	1075	18	AA235858	Human secreted pro
63	26	2.0	1076	21	AA235858	Human secreted pro
64	26	2.0	1076	21	AA235858	Human secreted pro
65	26	2.0	1108	24	AA235858	Human secreted pro
66	26	2.0	1108	24	AA235858	Human secreted pro
67	26	2.0	1126	21	AA235858	Human secreted pro
68	26	2.0	1144	22	AA235858	Human secreted pro
69	26	2.0	1144	22	AA235858	Human secreted pro
70	26	2.0	1145	22	AA235858	Human secreted pro
71	26	2.0	1149	22	AA235858	Human secreted pro
72	26	2.0	1185	21	AA235858	Human secreted pro
73	26	2.0	1256	21	AA235858	Human secreted pro
74	26	2.0	1263	20	AA235858	Human secreted pro
75	26	2.0	1334	21	AA235858	Human secreted pro
76	26	2.0	1331	20	AA235858	Human secreted pro
77	26	2.0	1331	21	AA235858	Human secreted pro
78	26	2.0	1331	21	AA235858	Human secreted pro
79	26	2.0	1334	20	AA235858	Human secreted pro
80	26	2.0	1350	21	AA235858	Human secreted pro
81	26	2.0	1350	21	AA235858	Human secreted pro
82	26	2.0	1350	21	AA235858	Human secreted pro

82	26	2.0	1352	20	AAx76333	Pneumocystis carin
83	26	2.0	1354	19	AAV59705	Human secreted pro
84	26	2.0	1354	24	ABs73699	Human CDNA #4 for
85	26	2.0	1433	20	AAz24885	Human secreted pro
86	26	2.0	1433	21	AAc69583	Human secreted pro
87	26	2.0	1458	21	AAc59929	Human secreted pro
88	26	2.0	1486	19	AAV59523	Human secreted pro
89	26	2.0	1486	24	ABs73510	Human CDNA #1 for
90	26	2.0	1574	23	AAs75598	DNA encoding novel
91	26	2.0	1576	21	AAz65287	Human secreted pro
92	26	2.0	1579	22	AAf28397	Nucleotide sequenc
93	26	2.0	1631	19	AAV59597	Human secreted pro
94	26	2.0	1631	24	ABs73535	Human CDNA #1 for
95	26	2.0	1636	21	AAz65285	Human secreted pro
96	26	2.0	1662	21	AAf22351	Human secreted pro
97	26	2.0	1662	25	ABz73597	Secreted protein-e
98	26	2.0	1662	25	ABz67192	Human secreted pro
99	26	2.0	1830	17	AAf07731	TNF-R p75ic-Bindin
100	26	2.0	1842	22	AAz60782	Human Cancer agent

ALIGNMENTS

RESULT 1

ID ABA91278 standard; cDNA; 1323 BP.

AC ABA91278;

DT 04-APR-2002 (first entry)

DE Cryptosporidium parvum antigenic polypeptide 2 (AG2) cDNA.

KW AG2; antigen; immunogen; cryptosporidiosis; infection; diagnosis;
KW therapy; vaccine; gene; ss.

OS *Cryptosporidium parvum*.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT / *tag= a

FT 3'UTR 591..1298

FT	polyA_signal	1054..105
----	--------------	-----------

FT	polyA_signal	1250...125

XX

XX

XXX

XX

XX

XX

XX

DR P-PSDB; AAM50658.

PT Novel isolated nucleic acid comprising coding sequence for an

PT vaccine components for treating or preventing C. parvum infection in

PT mammals -

XII

PS Claim 3; Fig 2A-B; 61pp; English.
XX

xx The present sequence comprises an incomplete open reading frame

CC for immunogenic CrypTosporidium parvum antigenic polypeptide 2
CC (A02, see AAWS0658). The sequence is a composite obtained from 5
CC overlapping cDNA clones that were isolated from a C. parvum
CC oocyst cDNA expression library by immunoscreening with human immune
CC serum. The open reading frame encodes a protein of 193 amino acids
CC with a predicted mol.wt. of 21.8 kDa and a pI of 6.22. The
CC invention relates to the discovery of agl and ag2 genes encoding
CC C. parvum antigenic polypeptides Ag1 and Ag2, the characterisation
CC of these proteins, and antibodies that recognise epitopes of Ag1
CC and Ag2. Also provided are recombinant vectors, host cells and
CC methods of producing recombinant C. parvum antigenic proteins.
CC The proteins, their immunogenic fragments and analogues, and/or
CC chimeric proteins, can be used either alone or in combination with
CC other C. parvum antigens in novel subunit vaccines to provide
CC protection from cryptosporidial infection in mammals. Antibodies
CC generated against the proteins can be used as therapeutic agents.
CC The antigens and antibodies can also be used as diagnostics of
CC infection.
XX
XX Sequence 1322BP; 481 A; 255 C; 206 G; 381 T; 0 other;

Query Match	100.0%;	Score 1323;	DB 24;	Length 1323;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1323;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	I	GAATGGGCA	CGAGATTTTTTTTTT	TCCTTTACCTATTCGA	TATAGTTCCTTGATTCAA	60
Db	1	GAATTCGGCACG	AGATTTTTTTTTTTTTTTT	CTTTTACCTATTTCAATTAAGTTAGTTCTTGATTCGA	60	
Qy	61	ACGATGCAAA	GTCAATATTTGTTTTTAATTCAGATGATCCGGAATTTTGAAAAACATTT	120		
Db	61	ACGATGCAAA	GTCAATATTTGTTTTTAATTCAGATGATCCGGAATTTTGAAAAACATTT	120		
Qy	121	CTACTAAATTCG	AAATTTAAATTTGAGCTTGCGCTTGATTAATGATGTTGGCTCGAGAGTG	180		
Db	121	CTACTAAATTCG	AAATTTAAATTTGAGCTTGCGCTTGATTAATGATGTTGGCTCGAGAGTG	180		
Qy	181	ATATTTTTTATC	CTTGATAGGAAACAGCCTCTTGAGAGCTGTAGATTAATCAATGCGCTGTG	240		
Db	181	ATATTTTTTATC	CTTGATAGGAAACAGCCTCTTGAGAGCTGTAGATTAATCAATGCGCTGTG	240		
Qy	241	TTTTCTATACAAA	ACAATGTTTTGAAAAGATGAAGACATTTGCTTTAAACCCCTTATATC	300		
Db	241	TTTTCTATACAAA	ACAATGTTTTGAAAAGATGAAGACATTTGCTTTAAACCCCTTATATC	300		
Qy	301	GCCTGAGAA	TAAAAATGACTTTTGCTTCGAGAAAGA	CTTAGCGACAATCTCCCAATCTT	360	
Db	301	GCCTGAGAA	TAAAAATGACTTTTGCTTCGAGAAAGA	CTTAGCGACAATCTCCCAATCTT	360	
Qy	361	CTAATTCGAA	TATATTCCTTTCTTCATCAGGAATTAATGCAATCTTTGGATATATTCGA	420		
Db	361	CTAATTCGAA	TATATTCCTTTCTTCATCAGGAATTAATGCAATCTTTGGATATATTCGA	420		
Qy	421	TTAACAAAT	TAAAAATTAACAACAATTCCTCTGGAATTCGCGGAATTCGAAATCCGCAATCCGA	480		
Db	421	TTAACAAAT	TAAAAATTAACAACAATTCCTCTGGAATTCGCGGAATTCGAAATCCGCAATCCGA	480		
Qy	481	CAATTGATAT	TTTGAGATATATGATGATATTTTCTTCTTTTAATTTAAACATGCGCCTA	540		
Db	481	CAATTGATAT	TTTGAGATATATGATGATATTTTCTTCTTTTAATTTAAACATGCGCCTA	540		
Qy	541	TAAAAGAAA	TGGCCACTTGATATATCAGAAATTCGAAATTCGAAATTTGAAATATATGAG	600		
Db	541	TAAAAGAAA	TGGCCACTTGATATATCAGAAATTCGAAATTTGAAATATATGAG	600		
Qy	601	CCACTAAATC	ACAATCCAGATCAGGTGGGAAATTAATTAATCTCAGAACCAAAAACACA	660		
Db	601	CCACTAAATC	ACAATCCAGATCAGGTGGGAAATTAATTAATCTCAGAACCAAAAACACA	660		
Qy	661	AAGCTCATC	AGTTGTAGACCGGATCCATACAGAAAGGCCCTCAAAATGATGAAATCA	720		
Db	661	AAGCTCATC	AGTTGTAGACCGGATCCATACAGAAAGGCCCTCAAAATGATGAAATCA	720		
Qy	720	AAAGCTCATC	AGTTGTAGACCGGATCCATACAGAAAGGCCCTCAAAATGATGAAATCA	720		
Db	720	AAAGCTCATC	AGTTGTAGACCGGATCCATACAGAAAGGCCCTCAAAATGATGAAATCA	720		

```

QY 721 GAATCTGTTGAGAGCAAGCCCTATTACTCACCAGAAAAAGAAATCAAGTTCAATCAT 780
DB 721 GAATCTGTTGAGAGCAAGCCCTATTACTCACCAGAAAAAGAAATCAAGTTCAATCAT 780
QY 781 CTTTCTGTGACAGTTCCAGATACCTGATCAAGTTCAATCTCTTCTGTACTATTCAGA 840
DB 781 CTTTCTGTGACAGTTCCAGATACCTGATCAAGTTCAATCTCTTCTGTACTATTCAGA 840
QY 841 GACTGTGACAGACTCAGATCAAGCCCTGTGACAAATTCAGAGACTGATCAAGTTCAA 900
DB 841 GACTGTGACAGACTCAGATCAAGCCCTGTGACAAATTCAGAGACTGATCAAGTTCAA 900
QY 901 TCATCTTCTGCTACTATACAGAAACAGATCCAGCTCAGATCACTCTGTAATCTTC 960
DB 901 TCATCTTCTGCTACTATACAGAAACAGATCCAGCTCAGATCACTCTGTAATCTTC 960
QY 961 CAGAGAAAGATTGAGACTCAGAAAGTTACCAATCTTCTACAGAACAACTCAAGCCA 1020
DB 961 CAGAGAAAGATTGAGACTCAGAAAGTTACCAATCTTCTACAGAACAACTCAAGCCA 1020
QY 1021 GCTACATATCTCTAACCAAGAAATGAAATCATATATATGAGAAAGTAATCGAGTTT 1080
DB 1021 GCTACATATCTCTAACCAAGAAATGAAATCATATATATGAGAAAGTAATCGAGTTT 1080
QY 1081 AATACACTAAATCTTCCAAATCAACCCCAATCTTTCAGCAAGCTGCGAGATGTGAAA 1140
DB 1081 AATACACTAAATCTTCCAAATCAACCCCAATCTTTCAGCAAGCTGCGAGATGTGAAA 1140
QY 1141 TATGGGAAAAAGATTAATATGTTGATGTGAGCAAGTAATCACTAAATATGACATTAT 1200
DB 1141 TATGGGAAAAAGATTAATATGTTGATGTGAGCAAGTAATCACTAAATATGACATTAT 1200
QY 1201 GAAGATCTCTCGAAAGAAATTTAGAAACAAATGTAAATGTGATTAATATATGAGC 1260
DB 1201 GAAGATCTCTCGAAAGAAATTTAGAAACAAATGTAAATGTGATTAATATATGAGC 1260
QY 1261 CTTAGCCATTTCCAAATATCTAAATTTGTCAACTCACTAAATCACTAAATCACTAA 1320
DB 1261 CTTAGCCATTTCCAAATATCTAAATTTGTCAACTCACTAAATCACTAAATCACTAA 1320
QY 1321 GAG 1323
DB 1321 GAG 1323

```

RESULT 2

AAV38762/c
ID AAV38762 standard; cDNA; 549 BP.

AAV38762;

18-SEP-1998 (first entry)

Complementary strand of cDNA nFSP18-549.

Serine protease inhibitor; SPI; anti-haematophagous ectoparasite;
KM vaccine; control; fleas; acarina; immune response; ss.

Ctenocephalides felis.

MO9820034-A2.

14-MAY-1998.

05-NOV-1997; 97WO-US20678.

07-NOV-1996; 96US-0745995.

(HESK-) HESKA CORP.

Brandt KS, Maddux JD, Silver GM, Wisniewski N;

WPI; 1998-286864/25.

```

XX New nucleic acid encoding new serine protease inhibitors from fleas
PT - and related antibodies, inhibitors, recombinant viruses and cells,
PT useful for control of fleas on animals, especially cats and dogs
PS Claim 2; Page 139; 176pp; English.
XX
XX The present sequence represents the complementary strand of nFSP18-549,
CC which encodes a Ctenocephalides felis serine protease inhibitor (SPI),
CC variable domain protein. The SPI protein represents a novel target for
CC anti-haematophagous ectoparasitic vaccines and drugs. Compositions
CC containing the inhibitor are used to control fleas (adults or larvae)
CC on mammals or birds, particularly cats and dogs. More generally the
CC composition is also used for control of insects and acarina, on animals
CC or on stored goods, plants, trees etc. including domestic pests and
CC insects that are vectors of disease. The SPI proteins are able to induce
CC a protective immune response against SPI.
XX
SQ Sequence 549 BP; 162 A; 103 C; 87 G; 197 T; 0 other;

```

Query Match 2.2%; Score 29; DB 19; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1295 AAGTAAAAAATTTAGAAACAAATGTAAATGTGATTAATATGAGC 1323
DB 51 AAGTAAAAAATTTAGAAACAAATGTAAATGTGATTAATATGAGC 23

RESULT 3

AAV38761
ID AAV38761 standard; cDNA; 549 BP.

AAV38761;

18-SEP-1998 (first entry)

cDNA nFSP18-549 encoding a SPI variable domain protein.

Serine protease inhibitor; SPI; anti-haematophagous ectoparasite;

KM vaccine; control; fleas; acarina; immune response; ss.

Ctenocephalides felis.

Key Location/Qualifiers

MO9820034-A2.

14-MAY-1998.

05-NOV-1997; 97WO-US20678.

07-NOV-1996; 96US-0745995.

(HESK-) HESKA CORP.

Brandt KS, Maddux JD, Silver GM, Wisniewski N;

WPI; 1998-286864/25.

P-PSDB; AAW62635.

New nucleic acid encoding new serine protease inhibitors from fleas

PT - and related antibodies, inhibitors, recombinant viruses and cells,

PT useful for control of fleas on animals, especially cats and dogs

PS Claim 2; Pages 137-136; 176pp; English.

The present sequence, designated nFSP18-549, encodes a Ctenocephalides

felis serine protease inhibitor (SPI) variable domain protein. The

complementary strand is represented by AAV38762. The SPI protein

represents a novel target for anti-haematophagous ectoparasitic vaccines

XX Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferris AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI Kyaw H, Lefleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX WPI, 1998-506364/43.
DR P-PSDB; AAW74778.
XX
PT New isolated human genes and the secreted polypeptide(s) they encode
PT -useful for diagnosis and treatment of e.g. cancer, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Claim 1, Page 280; 721pp; English.
XX
CC This sequence represents a nucleic acid molecule designated Gene 48 from
CC the human cDNA clone HFCAR174 (deposited as clone ATCC 97899 and ATCC
CC 209045) which encodes a secreted human protein. The gene can be used to
CC generate fusion proteins by linking to the gene to a human
CC immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of
CC the fused protein as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: AAV59511-V59612; amino acid sequences AAW74731-W75026)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Specific uses are described for each of the 186
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAV59511 for described uses).
XX
SO Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 other;
Query Match 2.2%; Score 29; DB 19; Length 1049;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
Db 1012 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1040

RESULT 5
ABS73545
ID ABS73545 standard; cDNA; 1049 BP.
XX
AC ABS73545;
XX
DT 15-JAN-2003 (first entry)
XX
DE Human cDNA #1 for novel secreted protein gene 48.
XX
KW Human; ss; gene; secreted protein; autoimmune disease; chemotaxis;
KW rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;
KW liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;
KW cerebrovascular disorder; cerebral ischemia; angiogenesis; sunburn;
KW nervous system disorders; Alzheimer's disease; infection; regeneration;
KW ocular disorder; corneal infection; wound healing; tissue regeneration;
KW epithelial cell proliferation; organ transplantation; food additive;
KW preservative; nutritional.
XX
OS Homo sapiens.
XX
PN US6420526-B1.
XX
PD 16-JUL-2002.
XX
PF 08-SEP-1998; 98US-0149476.
XX
PR 07-MAR-1997; 97US-038621P.
PR 07-MAR-1997; 97US-040161P.
PR 07-MAR-1997; 97US-040162P.
PR 07-MAR-1997; 97US-040163P.

PR 07-MAR-1997; 97US-040333P.
PR 07-MAR-1997; 97US-040334P.
PR 07-MAR-1997; 97US-040336P.
PR 07-MAR-1997; 97US-040626P.
PR 11-APR-1997; 97US-043311P.
PR 11-APR-1997; 97US-043312P.
PR 11-APR-1997; 97US-043313P.
PR 11-APR-1997; 97US-043314P.
PR 11-APR-1997; 97US-043315P.
PR 11-APR-1997; 97US-043368P.
PR 11-APR-1997; 97US-043369P.
PR 11-APR-1997; 97US-043376P.
PR 11-APR-1997; 97US-043378P.
PR 11-APR-1997; 97US-043380P.
PR 11-APR-1997; 97US-043669P.
PR 11-APR-1997; 97US-043670P.
PR 11-APR-1997; 97US-043671P.
PR 11-APR-1997; 97US-043672P.
PR 11-APR-1997; 97US-043674P.
PR 23-MAY-1997; 97US-047492P.
PR 23-MAY-1997; 97US-047500P.
PR 23-MAY-1997; 97US-047501P.
PR 23-MAY-1997; 97US-047502P.
PR 23-MAY-1997; 97US-047503P.
PR 23-MAY-1997; 97US-047581P.
PR 23-MAY-1997; 97US-047582P.
PR 23-MAY-1997; 97US-047583P.
PR 23-MAY-1997; 97US-047584P.
PR 23-MAY-1997; 97US-047585P.
PR 23-MAY-1997; 97US-047586P.
PR 23-MAY-1997; 97US-047587P.
PR 23-MAY-1997; 97US-047588P.
PR 23-MAY-1997; 97US-047589P.
PR 23-MAY-1997; 97US-047590P.
PR 23-MAY-1997; 97US-047592P.
PR 23-MAY-1997; 97US-047593P.
PR 23-MAY-1997; 97US-047594P.
PR 23-MAY-1997; 97US-047595P.
PR 23-MAY-1997; 97US-047596P.
PR 23-MAY-1997; 97US-047597P.
PR 23-MAY-1997; 97US-047598P.
PR 23-MAY-1997; 97US-047599P.
PR 23-MAY-1997; 97US-047600P.
PR 23-MAY-1997; 97US-047601P.
PR 23-MAY-1997; 97US-047612P.
PR 23-MAY-1997; 97US-047613P.
PR 23-MAY-1997; 97US-047614P.
PR 23-MAY-1997; 97US-047615P.
PR 23-MAY-1997; 97US-047617P.
PR 23-MAY-1997; 97US-047618P.
PR 23-MAY-1997; 97US-047632P.
PR 23-MAY-1997; 97US-047633P.
PR 06-JUN-1997; 97US-048964P.
PR 13-JUN-1997; 97US-048974P.
PR 08-JUL-1997; 97US-049610P.
PR 16-JUL-1997; 97US-051926P.
PR 18-AUG-1997; 97US-052874P.
PR 22-AUG-1997; 97US-055724P.
PR 22-AUG-1997; 97US-056530P.
PR 22-AUG-1997; 97US-056531P.
PR 22-AUG-1997; 97US-056532P.
PR 22-AUG-1997; 97US-056537P.
PR 22-AUG-1997; 97US-056562P.
PR 22-AUG-1997; 97US-056564P.
PR 22-AUG-1997; 97US-056564P.
PR 22-AUG-1997; 97US-056662P.
PR 22-AUG-1997; 97US-056664P.
PR 22-AUG-1997; 97US-056672P.
PR 22-AUG-1997; 97US-056674P.
PR 22-AUG-1997; 97US-056875P.
PR 22-AUG-1997; 97US-056876P.
PR 22-AUG-1997; 97US-056877P.

RESULT 11
ABAI083/C
ID ABAI083 standard; cDNA; 559 BP.
XX
AC ABAI083;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 3090.
XX
KW Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparthenogenic; antisklering; antianemic; antitubercic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antidiabetic; antidiabetic; antidiabetic; antidiabetic;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.
XX
OS Homo sapiens.
XX
PN W0200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001W0-US01334.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0235484.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0235837.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX MPI; 2001-541565/60.
XX P-PSDB; ABB17757.
XX
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases -
XX
XX Claim 1; SEQ ID NO 3090; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-AB21534) and proteins
XX (AB14678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 559 BP; 131 A; 120 C; 107 G; 197 T; 4 other;
SQ
Query Match 2.0%; Score 27; DB 22; Length 559;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1295 AAGTAAATATATATATATATATATCTCG 1321
DB 29 AAGTAAATATATATATATATATATATCTCG 3
RESULT 12
AAS35580/C
ID AAS35580 standard; CDNA; 708 BP.
XX
XX AAS35580;
XX
XX 17-DEC-2001 (first entry)
XX
XX Human cardiovascular system antigen cDNA polynucleotide SEQ ID No 465.
XX
XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;

KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
KW antineumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nocrotrophic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerrary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.
XX
XX Homo sapiens.
XX
XX WO200155321-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01340.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218280.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226688.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.
XX 14-SEP-2000; 2000US-0232397.
XX 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234224.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 25-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 01-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249247.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251899.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-451930/48.
 DR P-PSDB; AAU22306.
 PT
 PT New cardiovascular system related polynucleotides and polypeptides,
 PT useful for diagnosing, treating and/or preventing disorders of the
 PT cardiovascular system -
 XX
 PS
 PS Claim 1; SEQ ID NO 465; 674pp; English.
 XX
 CC Sequences AAS35126-AAS35740 represent cDNA molecules, which encode the
 CC Cardiovascular system antigen polypeptides of the invention.
 CC Cardiovascular system antigens and their associated polynucleotides are
 CC useful in the diagnosis, treatment and prevention of various types of
 CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
 CC chickens or sheep. A pathological condition can be determined by
 CC detecting the presence or absence of a mutation in a cardiovascular
 CC system antigen polynucleotide. The treatable disorders include autoimmune
 CC diseases such as rheumatoid arthritis, hyperproliferative disorders such
 CC as neoplasms of the breast or liver, cardiovascular disorders such as
 CC cardiac arrest, cerebrovascular disorders such as cerebral ischemia,
 CC nervous system disorders such as Alzheimer's disease, infections caused
 CC by bacteria, viruses and fungi, ocular disorders such as corneal
 CC infection, endocrine disorders such as premature labour and infertility,
 CC gastrointestinal disorders such as Crohn's disease, renal disorders such
 CC as glomerulonephritis and respiratory disorders such as asthma and
 CC pleurisy. The polypeptides can also be used to aid wound healing, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, to regenerate tissues and in chemotaxis.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Query Match 2.0%; Score 27; DB 22; Length 708;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAACTCGAG 1323
 Db 51 GTAAAAAAAAAAAAAAAACTCGAG 25

RESULT 13
 AAA01696
 ID AAA01696 standard; cDNA; 740 BP.
 XX
 AC AAA01696;
 XX
 DT 19-MAY-2000 (first entry)
 XX
 DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:1687.
 XX
 XX Human; colon cancer; tumour; diagnosis; gene expression product;
 KM probe; detection; cancerous state; metastasis; identification;
 KM breast cancer; oestrogen receptor-positive breast cancer; therapy;
 KM oestrogen receptor-negative breast cancer; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 XX MO958675-A2.
 PN

```

XX 18-NOV-1999.
PD
XX
XX 13-MAY-1999; 99WO-US10602.
PF
XX 14-MAY-1998; 98US-0085426.
PR 15-MAY-1998; 98US-0085537.
PR 15-MAY-1998; 98US-0085696.
PR 21-OCT-1998; 98US-0105234.
PR 27-OCT-1998; 98US-0105877.
XX
XX (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kasam A;
PI Lamson G, Dmanac R, Ckrvenjakov R, Dickson M, Dmanac S, Labat I;
PI Leebkowitz D, Kita D, Garcia V, Jones LM, Stache-Crain B;
XX
XX WPI; 2000-126369/11.
DR
XX
XX Polynucleotide library used to determine cancerous states of mammalian
PT cells -
XX
XX Claim 1; Page 613; 1097P; English.
PS
XX
XX AAA0010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotides sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC pre-metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.
CC
XX
XX Sequence 740 BP; 247 A; 110 C; 134 G; 243 T; 6 other;
SQ
Query Match 2.0%; Score 27; DB 21; Length 740;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
Db 672 GTAAAAAAAAAAAAAAAAAACTCGAG 698
RESULT 14
AAT63042
ID AAT63042 standard; cDNA; 1017 BP.
XX
XX AAT63042;
AC
XX
XX 22-JUL-1997 (first entry)
DT
XX
XX Aspergillus niger cellulase cDNA.
DE
XX
XX Cellulase; protein detection; enzyme detection; library screening;
KW ds.
XX
XX Aspergillus niger strain N400 (CBS120.49).
OS
XX
XX Key Location/Qualifiers
FH 57..776
FT CDS
FT /*tag= a

```

```

XX
XX WO9713853-A2.
PN
XX
XX 17-APR-1997.
PD
XX
XX 14-OCT-1996; 96WO-EP04510.
PF
XX 13-OCT-1995; 95EP-0202777.
PR
XX (KONN ) GIST-BROCADES BV.
PA
XX
XX De Graaff LH, Van Den Broeck HC, Van Ooyen AJJ, Visser J;
PI
XX
XX WPI; 1997-235889/21.
DR
XX
XX P-PSDB; AAM14595.
DR
XX
XX Identifying DNA fragments encoding proteins, e.g. for new enzyme
PT discovery - by direct screening of a cDNA library in bacteria
PT transformed with DNA from eukaryotic organism producing the protein
PT
XX
XX Claim 5; Page 13-14; 30pp; English.
PS
XX
XX A cDNA clone (AAT63042), deposited as CBS 589.95, codes for a
CC cellulase (AAM14595) of Aspergillus niger N400 (CBS120.49). It was
CC isolated using a method for identifying a DNA fragment encoding a
CC protein of interest. The method involves the direct screening of a
CC cDNA library prep'd. in bacteria (e.g. E. coli) transformed with DNA
CC from a eukaryotic organism (e.g. A. niger) that produces the
CC protein. Screening for cellulase-producing clones can be performed
CC on plates containing CM-cellulase and Congo Red. The method was
CC utilised in the identification of cDNA clones (AAT63042-46) coding
CC for A. niger cellulase, xylanase and arabinoxylan degrading enzymes
CC (AAM14595-99).
XX
XX
XX Sequence 1017 BP; 258 A; 267 C; 265 G; 227 T; 0 other;
SQ
Query Match 2.0%; Score 27; DB 18; Length 1017;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
Db 991 GTAAAAAAAAAAAAAAAAAACTCGAG 1017
RESULT 15
AAT60426
ID AAT60426 standard; cDNA; 1017 BP.
XX
XX AAT60426;
AC
XX
XX 08-JUL-1997 (first entry)
DT
XX
XX Cellulase cDNA clone 12.
DE
XX
XX Cellulase; carboxymethylcellulase; endoglucanase; beta-glucanase;
KW cellulolytic enzyme; wine; beer; juice clarification; paper; pulp;
KW textile; feedstuff; ds.
XX
XX Aspergillus niger strain N400 (CBS120.49).
OS
XX
XX Key Location/Qualifiers
FH 57..776
FT CDS
FT /*tag= a
XX
XX WO9713862-A1.
PN
XX
XX 17-APR-1997.
PD
XX
XX 14-OCT-1996; 96WO-EP04511.
PF
XX 13-OCT-1995; 95EP-0202776.
PR
XX

```

PA (KONN) GIST-BROCADES BV.
 XX
 XX De Graaf LH, Van Den Broeck HC, Van Ooyen AJJ, Visser J;
 PI WPI; 1997-235898/21.
 XX P-PSDB; AAW15562.
 DR A6pexgluc cellololytic enzyme - degrades carboxymethyl cellulose,
 XX endoglucan and beta-glucan, used e.g. in production of wine, animal
 PT feeds etc.
 CC Claim 2; Page 15-16; 28pp; English.
 XX
 XX CC DNA clones 12 (AA760426) and 64 (AA760427) are respectively deposited
 CC as CBS 589.95 and CBS 662.95 and code for enzymes (AAW15562 and
 CC AAW15563) that have carboxymethylcellulase, endoglucanase and beta-
 CC glucanase activities. They were obtd. by screening colonies of an
 CC A. niger cDNA library, prepd. in E. coli using a plasmid vector, in
 CC a plate assay using CM-cellulase overlay with Congo red staining.
 CC The cDNA clones can be expressed in yeast, bacteria or fungi
 CC for use in degrading cellulose and glycans, esp. in the manufacture
 CC of wine, beer or fruit/vegetable juices (e.g. to liquify plant walls
 CC and reduce membrane fouling) and in the textile, detergent and
 CC paper/pulp industries e.g. to impart a worn appearance to clothing.
 CC The cellulases can also be used to improve animal feed utilisation.
 XX
 SO Sequence 1017 BP; 258 A; 267 C; 265 G; 227 T; 0 other;
 Query Match 2.0%; Score 27; DB 18; Length 1017;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
 DB 991 GTAAAAAAAAAAAAAAAAAACTCGAG 1017
 RESULT 16
 AA207191.
 ID AA207191 standard; cDNA; 1061 BP.
 XX
 AC AA207191;
 XX
 DT 13-OCT-1999 (first entry)
 XX
 DE Human lung tumour protein SAL-3 5' cDNA sequence.
 XX
 XX Human lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
 KM immunotherapy; detection; inhibition; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO938973-A2.
 XX
 PD 05-AUG-1999.
 XX
 PF 26-JAN-1999; 99WO-US01642.
 XX
 XX 22-DEC-1998; 98US-0219245.
 PR 28-JAN-1998; 98US-0015022.
 PR 28-JAN-1998; 98US-0015029.
 PR 18-MAR-1998; 98US-0040828.
 PR 18-MAR-1998; 98US-0040831.
 PR 23-JUL-1998; 98US-0122191.
 PR 23-JUL-1998; 98US-0122192.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Frudakis TN, Lodes MJ, Mohamath R, Reed SG;
 XX WPI; 1999-479187/40.
 DR P-PSDB; AAY29510.
 XX

PT Lung tumour specific polynucleotides for inhibiting the development
 PT of lung cancer
 XX
 XX Claim 1; Page 132; 171pp; English.
 PS
 XX The present invention describes lung tumour specific polynucleotides
 CC and tumour antigens. AA207144 to AA207246 and AA208301 to AA208325
 CC represent specifically claimed polynucleotides, and AAY29486 to AAY29571
 CC represent amino acid sequences from the present invention. The lung
 CC tumour specific polynucleotides and polypeptides can be used in
 CC pharmaceutical compositions and vaccines to inhibit the development of
 CC lung cancer. They can also be used to detect lung cancer in a patient.
 CC Probes and antibodies derived from the lung tumour sequences are useful
 CC in detection of lung cancer.
 XX
 SO Sequence 1061 BP; 219 A; 294 C; 350 G; 198 T; 0 other;
 Query Match 2.0%; Score 27; DB 20; Length 1061;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
 DB 1035 GTAAAAAAAAAAAAAAAAAACTCGAG 1061
 RESULT 17
 AAC79143
 ID AAC79143 standard; cDNA; 1061 BP.
 XX
 AC AAC79143;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Human lung tumour-specific cDNA #96.
 XX
 XX Lung tumour protein; lung cancer; cytostatic; vaccine; ss.
 KM
 XX Homo sapiens.
 OS
 PN WO200060077-A2.
 XX
 PD 12-OCT-2000.
 XX
 PF 30-MAR-2000; 2000WO-US08560.
 XX
 XX 02-APR-1999; 99US-0285323.
 PR 09-AUG-1999; 99US-0370838.
 PR 30-DEC-1999; 99US-0476235.
 PR 03-MAR-2000; 2000US-0518809.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Lodes MJ, Mohamath R, Secrist H;
 XX WPI; 2000-638466/61.
 DR P-PSDB; AAB44458.
 XX
 XX Novel lung tumor polypeptides and polynucleotides, useful for
 PT detecting, monitoring or treating cancer, especially lung cancer -
 XX
 PS Claim 1; Page 148; 243pp; English.
 XX
 XX The present sequence is given in a specification relating to compounds
 CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
 CC least an immunogenic part of a lung tumour protein are disclosed.
 CC The polypeptides are useful for inhibiting the development of cancer,
 CC especially lung cancer. Samples of T cells expressing the polypeptides
 CC may be used to inhibit the development of cancer. The polypeptides are
 CC also useful for detecting and monitoring the progression of cancer,
 CC especially lung cancer.
 XX
 SO Sequence 1061 BP; 219 A; 294 C; 350 G; 198 T; 0 other;

Query Match 2.0%; Score 27; DB 21; Length 1061;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
 |||||
 DB 1035 GTAAAAAAAAAAAAAAAAAACTCGAG 1061

RESULT 18

AAD23218
 ID AAD23218 standard; cDNA; 1061 BP.

AC AAD23218;

DT 26-FEB-2002 (first entry)

DE Human lung tumour-specific protein SAL-3 cDNA.

XX Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
 KW antisense-therapy; vaccine; immune response; lung cancer; SAL-3; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH 15..644
 FT /*tag= a
 FT /product= "Human SAL-3 protein"
 FT /note= "CDS does not include start codon"
 FT /partial

XX MO200172295-A2.

XX 04-OCT-2001.

XX 28-MAR-2001; 2001WO-US09991.

XX 29-MAR-2000; 2000US-0538037.

XX 05-JUN-2000; 2000US-0588937.

XX 18-AUG-2000; 2000US-0640878.

XX 22-SEP-2000; 2000US-234517P.

XX 01-NOV-2000; 2000US-0704512.

XX 14-DEC-2000; 2000US-0738973.

XX (CORI-) CORIXA CORP.

XX Read SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;
 PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;

XX WPI; 2001-639201/73.
 DR P-PSDB; AAE13799.

XX New human lung-specific polynucleotides and polypeptides for the
 PT diagnosis and treatment of disease e.g. lung cancer -

XX Example 4; Page 201-202; 378pp; English.

XX The invention relates to isolated lung tumour-specific proteins and
 CC their corresponding cDNA molecules. Lung tumour-specific proteins and
 CC their antigen-presenting cells are useful for stimulating and/or
 CC expanding T cells specific for a tumour protein, and for inhibiting
 CC the development of cancer. The invention also relates to a composition
 CC useful for stimulating an immune response, and for treating cancer. The
 CC lung tumour specific oligonucleotide is useful in gene therapy and for
 CC diagnosis, detection and treatment of lung cancer. The present sequence
 CC is a cDNA encoding human lung tumour-specific protein.

XX Sequence 1061 BP; 219 A; 294 C; 350 G; 198 T; 0 other;

XX Query Match 2.0%; Score 27; DB 23; Length 1061;

XX Best Local Similarity 100.0%; Pred. No. 2.1;
 XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
 |||||
 DB 1035 GTAAAAAAAAAAAAAAAAAACTCGAG 1061

RESULT 19

AA215856
 ID AA215856 standard; cDNA; 1249 BP.

AC AA215856;

DT 12-OCT-1999 (first entry)

DE Human gene expression product cDNA sequence SEQ ID NO:3325.

XX Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

OS Homo sapiens.

XX MO9938972-A2.

XX 05-AUG-1999.

XX 28-JAN-1999; 99WO-US01619.

XX 03-APR-1998; 98US-0080666.

XX 28-JAN-1998; 98US-0072910.

XX 24-FEB-1998; 98US-0075954.

XX 31-MAR-1998; 98US-0080114.

XX 03-APR-1998; 98US-0080515.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lamsom G, Leschowitz D, Pot D, Randazzo F, Reinhard C;
 PI Schache-Crain B, Suduth-Klinger J, Williams LT;

XX WPI; 1999-494092/41.

XX Novel human genes and their expression products which are
 PT differentially expressed in different cell types

XX Claim 1; Page 1593; 2479pp; English.

XX The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AA212532 to AA217779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The
 CC polynucleotides can be used as a source of primers and probes, which can
 CC be used for a variety of purpose, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensics, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and
 CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels between
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
 CC identify a genetic predisposition or susceptibility to a disease such as
 CC cancer). The polynucleotides of the invention are especially used in the
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
 CC and lung cancer. The polynucleotides can also be used to screen for
 CC peptide analogues and antagonists.

XX Sequence 1249 BP; 426 A; 160 C; 223 G; 396 T; 44 other;

Query Match 2.0%; Score 27; DB 20; Length 1249;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1323

Db 1201 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1227

RESULT 20

AA565294
ID AA565294 standard; cDNA; 1414 BP.

AC AA565294;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #1098.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PsDB; ABBG1107.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnosis; forensic; gene mapping; identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

PS Claim 1; SEQ ID NO 1098; 103bp; English.

XX The invention relates to isolated polynucleotide (II) and

XX polypeptide (II) sequences. (I) is useful as hybridization probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensic, gene mapping, identification of mutations in

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. AA564197-AA594564 represent novel human

XX diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

CC

XX

SO Sequence 1414 BP; 362 A; 343 C; 358 G; 351 T; 0 other;

Query Match 2.0%; Score 27; DB 23; Length 1414;

Best Local Similarity 100.0%; Pred. No. 2;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1323

Db 1372 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1398

RESULT 21

ABL90278
ID ABL90278 standard; cDNA; 1427 BP.

AC ABL90278;

DT 24-MAY-2002 (first entry)

DE Human polynucleotide SEQ ID NO 840.

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

KM anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anticancer;

KM vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;

KM cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;

KM neurological disease; infection; human; secreted protein; gene; ss.

OS Homo sapiens.

PN WO200190304-A2.

PD 29-NOV-2001.

PF 18-MAY-2001; 2001WO-US16450.

PR 19-MAY-2000; 2000US-205515P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

DR WPI; 2002-122018/16.

DR P-PsDB; ABB89869.

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and

PT prevention of neural, immune system, muscular, reproductive,

PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders -

PS Claim 4; SEQ ID NO 840; 2081bp + Sequence Listing; English.

XX The invention relates to novel genes (ABL9449-ABL90853) and proteins

XX (ABL89040-ABL90444) useful for preventing, treating or ameliorating

XX medical conditions e.g. by protein or gene therapy. The genes are

XX isolated from a range of human tissues disclosed in the specification.

XX The nucleic acids, proteins, antibodies and (ant)agonists are useful

XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast

XX and ovarian cancer and other cancers of the adrenal gland, bone, bone

XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune

XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative

XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;

XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and

XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal

XX and parasitic infections.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

CC

XX

SO Sequence 1427 BP; 255 A; 472 C; 440 G; 253 T; 7 other;

Query Match 2.0%; Score 27; DB 24; Length 1427;

Best Local Similarity 100.0%; Pred. No. 2;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAACTCGAG 1323
Db 1199 GTAAAAAAAAAAAAAAAACTCGAG 1225

RESULT 22

AAC77911
ID AAC77911 standard; cDNA; 1553 BP.

AC AAC77911;

DT 08-FEB-2001 (first entry)

DE Human cancer associated gene sequence SEQ ID NO:305.

XX Human, cancer associated gene; cancer antigen, detection; cancer;
XX diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
XX antidiabetic; antisthmatic; antirheumatic; antiarthritic; antiviral;
XX antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;
XX dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
XX vasotropic; antipsoriatic; angiogenic; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX allergic reaction; graft versus host disease; organ rejection;
XX haemostatic; thrombolytic; cardiovascular disorder; infection;
XX neurological disease; drug screening; ss.

OS Homo sapiens.

XX W0200055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05882.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

XX P-PSDB; AAB43702.

XX Novel isolated nucleic acids comprising sequences encoding peptides

XX useful for treating or diagnosing e.g. cancer -

XX Claim 1; Page 863-864; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given
XX in AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnery; immunomodulator;
XX antidiabetic; antisthmatic; antirheumatic; antiarthritic;
XX antiinflammatory; antihypertoid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
XX nootropic; vasotropic; antipsoriatic and angiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention.

XX Sequence 1553 BP; 451 A; 351 C; 346 G; 403 T; 2 other;

XX Query Match 2.0%; Score 27; DB 21; Length 1553;

Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAACTCGAG 1323
Db 1527 GTAAAAAAAAAAAAAAAACTCGAG 1553

RESULT 23

AAC93515
ID AAC93515 standard; cDNA; 2509 BP.

AC AAC93515;

DT 16-FEB-2001 (first entry)

DE Human secreted protein gene 37 SEQ ID NO:47.

XX Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
XX antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
XX nootropic; neuroprotective; antibacterial; virucide; fungicide;
XX ophthalmological; vulnery; autoimmune disease; rheumatoid arthritis;
XX hyperproliferative disorders; cancer; cardiovascular disorder;
XX cardiac arrest; cerebrovascular disorder; nervous system disorder;
XX Alzheimer's disease; ocular disorder; wound healing; skin aging; ss.

OS Homo sapiens.

XX W0200061626-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US09066.

XX 09-APR-1999; 99US-0128698.

XX 20-JAN-2000; 2000US-0176926.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-619227/59.

XX P-PSDB; AAB51863.

XX New nucleic acid molecules encoding 49 human secreted proteins for

XX diagnosing, preventing or ameliorating medical conditions and used for

XX food additives or preservatives -

XX Claim 1; Page 454-455; 516pp; English.

XX Polynucleotide sequences AAC93479 - AAC93527 represent cDNA encoding
XX human secreted proteins AAB51827 - AAB51875. Sequences AAB51876 -
XX AAB51927 represent alternative polypeptides encoded by the genes, and
XX amino acid sequences with which they share homology. The genes and
XX proteins have activities dependent on the tissues and cells in which they
XX are expressed. Examples of their activities include immunosuppressive;
XX antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
XX vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
XX virucide; fungicide; ophthalmological; and vulnery. The secreted
XX proteins, polynucleotides, antagonists and agonists may be useful in
XX treating, preventing and/or diagnosing diseases and disorders such as
XX autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease,
XX infections caused by bacteria, viruses and fungi and ocular disorders
XX e.g. corneal infection. The polypeptides can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues and in chemotaxis. The
XX polypeptides can also be used as a food additive or preservative to
XX increase or decrease storage capabilities, fat content, lipid, protein,

CC carbohydrate, vitamins, minerals, cofactors and other nutritional
CC components. Oligonucleotides AAC93470 - AAC93478 and peptide AAB51826 are
CC used in the isolation and characterisation of the proteins and
CC polynucleotides of the invention.

SO Sequence 2509 BP; 635 A; 568 C; 689 G; 614 T; 3 other;

Query Match 2.0%; Score 27; DB 21; Length 2509;

Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAACTCGAG 1323

Db 2483 GTAAAAAAAAAAAAAAAACTCGAG 2509

RESULT 24

ID AAC60009 standard; cDNA; 2528 BP.

XX AAC60009;

XX 26-JAN-2001 (first entry)

DE Human secreted protein gene 44 SEQ ID NO:54.

XX Human; secreted protein; diagnosis; cytosolic; immunosuppressive;
XX neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antineoplastic; antitumor; antiparasitic; anticonvulsant;
XX antidiabetic; antifungal; antiparasitic; cardiac; gene therapy;
XX cancer; immune disorder; cardiovascular disorder; wound healing;
XX neurological disease; infectious disease; chromosome identification; ss.

OS Homo sapiens.

XX WO200058356-A1.

XX 05-OCT-2000.

PD 22-MAR-2000; 2000MO-US07535.

XX 26-MAR-1999; 99US-0126511.

PR 17-DEC-1999; 99US-0172413.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-594639/56.

DR P-PSDB; AAB34816.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer; immune disorders,
PT cardiovascular disorders and neurological diseases -

XX Claim 1; Page 371; 425pp; English.

XX The polynucleotide sequences given in AAC59966 to AAC60015 encode the
XX human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
XX AAB34852 represent human secreted polypeptide sequences and proteins
XX homologous to them, which are given in the exemplification of the present
XX invention. Human secreted proteins have activities based on the tissues
XX and cells the genes are expressed in. Examples of activities include:
XX cytotoxic; immunosuppressive; neurotropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antineoplastic; antitumor;
XX antidiabetic; anticonvulsant; antibacterial; antifungal; antiparasitic; and
XX cardiant. The polynucleotides and polypeptides are useful for
XX preventing, treating or ameliorating a medical condition in e.g. humans,
XX mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
XX polypeptides can also be used as a food additive or preservative to
XX increase or decrease storage capabilities. The polynucleotides are
XX useful for chromosome identification. They are also useful as probes for
XX diagnosing a disorder related to the female reproductive system,

CC particularly breast and/or ovary cancer. They are also useful in the gene
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC agonists and antagonists from the present invention are useful in the
CC diagnosis, treatment and prevention of cancer, immune disorders,
CC cardiovascular disorders, wound healing, neurological diseases and
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
CC used in the exemplification of the present invention.

SO Sequence 2528 BP; 772 A; 428 C; 545 G; 781 T; 2 other;

Query Match 2.0%; Score 27; DB 21; Length 2528;

Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1297 GTAAAAAAAAAAAAAAAACTCGAG 1323

Db 2462 GTAAAAAAAAAAAAAAAACTCGAG 2488

RESULT 25

ID ABL01572 standard; cDNA; 2785 BP.

XX ABL01572;

XX 15-MAR-2002 (first entry)

DE Human secreted protein encoding polynucleotide SEQ ID NO 17.

XX Human; neurotropic; neuroprotective; cytosolic; dermatological; virucide;
XX immunosuppressive; antineoplastic; anti-HIV; antibacterial; vulnery;
XX antiparasitic; antidiabetic; antitumor; antiparasitic; cancer;
XX antineoplastic; hepatotropic; cerebroprotective; antineoplastic;
XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX antiparasitic; cardiac; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.

OS Homo sapiens.

XX WO200183510-A1.

XX 08-NOV-2001.

PD 26-APR-2001; 2001MO-US13318.

XX 02-MAY-2000; 2000US-201194P.

PR 16-JUN-2000; 2000US-212142P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Komatsoulis G, Ruben SM, Rosen CA;

XX WPI; 2002-121886/16.

DR P-PSDB; ABB57400.

XX An isolated nucleic acid molecule encoding a human secreted protein
PT (SP) is useful in preventing, treating or ameliorating a disorder e.g.,
PT Alzheimer's disease and cancers -

XX Claim 1; Page 437-438; 496pp; English.

XX The invention relates to novel genes (ABL01566-ABL01594) and proteins
XX (ABB57394-ABB57456) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;

PS Claim 4; Page 191-192; 251pp; English

DR WPI; 1999-479187/40.

DR WPI; 1999-479187/40.

XX Lung tumour specific polynucleotides for inhibiting the development
 PT of lung cancer
 XX
 XX Claim 13, Page 62-63, 171pp, English.
 PS
 CC The present invention describes lung tumour specific polynucleotides
 CC and tumour antigens. AA207144 to AA207246 and AA208301 to AA208325
 CC represent specifically claimed polynucleotides, and AA29486 to AA29571
 CC represent amino acid sequences from the present invention. The lung
 CC tumour specific polynucleotides and polypeptides can be used in
 CC pharmaceutical compositions and vaccines to inhibit the development of
 CC lung cancer. They can also be used to detect lung cancer in a patient.
 CC Probes and antibodies derived from the lung tumour sequences are useful
 CC in detection of lung cancer.
 CC
 SO Sequence 317 BP, 111 A, 55 C, 78 G, 73 T, 0 other;
 Query Match 2.0%; Score 26; DB 20; Length 317;
 Best Local Similarity 100.0%; Pred. No. 5.7;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
 Db 292 TAAAAAAAAAAAAAAAAAATCGAG 317
 RESULT 28
 AAC79062
 ID AAC79062 standard; CDNA; 317 BP.
 XX
 AC AAC79062;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Human lung tumour-specific cDNA #17.
 DE
 XX Lung tumour protein; lung cancer; cytostatic; vaccine; ss.
 KM
 XX Homo sapiens.
 OS
 XX MO200060077-A2.
 PN
 PD 12-OCT-2000.
 XX
 PF 30-MAR-2000; 2000MO-US08560.
 XX
 PR 02-APR-1999; 99US-0285323.
 PR 09-AUG-1999; 99US-0370838.
 PR 30-DEC-1999; 99US-0476235.
 PR 03-MAR-2000; 2000US-0518809.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Lodes MJ, Mohamath R, Secrist H;
 PI
 DR WPI; 2000-638466/61.
 DR P-PSDB; AAB44407.
 XX
 PT Novel lung tumor polypeptides and polynucleotides, useful for
 PT detecting, monitoring or treating cancer, especially lung cancer -
 XX
 PS Example 2, Page 90; 243pp, English.
 XX
 CC The present sequence is given in a specification relating to compounds
 CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
 CC least an immunogenic part of a lung tumour protein are disclosed.
 CC The polypeptides are useful for inhibiting the development of cancer.
 CC especially lung cancer. Samples of T cells expressing the polypeptides
 CC may be used to inhibit the development of cancer. The polypeptides are
 CC also useful for detecting and monitoring the progression of cancer,
 CC especially lung cancer.
 CC

SO Sequence 317 BP, 111 A, 55 C, 78 G, 73 T, 0 other;
 Query Match 2.0%; Score 26; DB 21; Length 317;
 Best Local Similarity 100.0%; Pred. No. 5.7;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
 Db 292 TAAAAAAAAAAAAAAAAAATCGAG 317
 RESULT 29
 AAC23137
 ID AAC23137 standard; CDNA; 317 BP.
 XX
 AC AAC23137;
 XX
 DT 26-FEB-2002 (first entry)
 XX
 DE Human lung tumour-specific protein LT86-1 cDNA.
 DE
 XX Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
 KM anticancer-therapy; vaccine; immune response; lung cancer; LT86-1; ss.
 XX
 OS Homo sapiens.
 OS
 FH Homo sapiens.
 FH
 FT Key Location/Qualifiers
 FT CDS 1..240
 FT /tag= a
 FT /product= "Human LT86-1 protein"
 FT /note= "CDS does not include start and stop codon"
 FT /partial
 XX
 PN MO200172295-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 28-MAR-2001; 2001MO-US09991.
 XX
 PR 29-MAR-2000; 2000US-0538037.
 PR 05-JUN-2000; 2000US-0588937.
 PR 18-AUG-2000; 2000US-0640878.
 PR 22-SEP-2000; 2000US-234517P.
 PR 01-NOV-2000; 2000US-0704512.
 PR 14-DEC-2000; 2000US-0738973.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY,
 PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
 XX
 DR WPI; 2001-639201/73.
 DR P-PSDB; AAE13748.
 XX
 PT New human lung-specific polynucleotides and polypeptides for the
 PT diagnosis and treatment of disease e.g. lung cancer -
 XX
 PS Example 2, Page 147; 378pp, English.
 XX
 CC The invention relates to isolated lung tumour-specific proteins and
 CC their corresponding cDNA molecules. Lung tumour-specific proteins and
 CC their antigen-presenting cells are useful for stimulating and/or
 CC expanding T cells specific for a tumour protein, and for inhibiting
 CC the development of cancer. The invention also relates to a composition
 CC useful for stimulating an immune response, and for treating cancer. The
 CC lung tumour specific oligonucleotide is useful in gene therapy and for
 CC diagnosis, detection and treatment of lung cancer. The present sequence
 CC is a cDNA encoding human lung tumour-specific protein.
 CC
 SO Sequence 317 BP, 111 A, 55 C, 78 G, 73 T, 0 other;
 Query Match 2.0%; Score 26; DB 23; Length 317;
 Best Local Similarity 100.0%; Pred. No. 5.7;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
 Db 292 TAAAAAAAAAAAAAAAACTCGAG 317

RESULT 30
 AAC81737
 ID AAC81737 standard; DNA; 352 BP.

AC AAC81737;
 DT 26-FEB-2001 (first entry)
 XX

DE Human secreted protein coding sequence SEQ ID NO: 38.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW antineoplastic; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; ss.

XX Homo sapiens.
 OS
 PN MO200061627-A1.
 XX
 PD 19-OCT-2000.

PF 06-APR-2000; 2000MO-US09067.

PR 09-APR-1999; 99US-0128697.
 PR 20-JAN-2000; 2000US-0176929.

XX (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.

PI Ruben SM, Komatsoulis G;

DR WPI: 2000-647419/62.
 DR P-PSDB; AAB45413.

PT New nucleic acid molecules encoding 49 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -

PS Claim 1; Page 444; 521pp; English.

XX The invention relates to the isolation of genes AAC81710-C81758 encoding
 CC 49 human secreted proteins AAB45386-B54534. The genes can be used to
 CC generate fusion proteins by linking to the gene for the human
 CC immunoglobulin G Fc portion (SE01D1) for increasing the stability of
 CC the fusion protein as compared to the human protein only. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated
 CC from a range of human tissues disclosed in the specification. The
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.

XX Sequence 352 BP; 79 A; 89 C; 123 G; 59 T; 2 other;

Query Match 2.0%; Score 26; DB 21; Length 352;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
 Db 327 TAAAAAAAAAAAAAAAACTCGAG 352

RESULT 31
 AAC95544
 ID AAC95544 standard; cDNA; 385 BP.

AC AAC95544;
 DT 21-FEB-2001 (first entry)
 XX

DE Human secreted protein gene 24 SEQ ID NO:34.

XX Human, secreted protein; cytosolic; immunosuppressive; nootropic;
 KW neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
 KW antineoplastic; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; cancer; immune disease; allergy;
 KW cardiovascular disorder; wound healing; infection; neurological disease;
 KW ss.

XX Homo sapiens.
 OS
 PN MO200061596-A1.
 XX
 PD 19-OCT-2000.

PF 06-APR-2000; 2000MO-US08983.

PR 09-APR-1999; 99US-0128703.
 PR 20-JAN-2000; 2000US-0176068.

XX (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.

PI Ruben SM, Komatsoulis G;

DR WPI: 2000-611865/58.
 DR P-PSDB; AAB52035.

PT Fifty nucleic acid molecules encoding human secreted proteins, useful
 PT in the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases -

PS Claim 1; Page 429; 505pp; English.

XX Polynucleotide sequences AAC95521 - AAC95570 represent cDNA encoding 50
 CC human secreted proteins AAB52012 - AAB52161. Sequences AAB52062 -
 CC AAB52103 represent alternative polypeptides encoded by the genes, and
 CC amino acid sequences with which they share homology. The genes and
 CC proteins have activities dependent on the tissues and cells in which they
 CC are expressed. Examples of their activities include cytosolic; antiallergic;
 CC immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic;
 CC hepatotropic; antidiabetic; antiinflammatory; antitumor; vulnery;
 CC anticonvulsant; antibacterial; antifungal; antiparasitic; and cardiant.
 CC The secreted proteins, polypeptides, antagonists and agonists may be
 CC useful in treating, preventing and/or diagnosing diseases and disorders
 CC such as cancer, particularly breast and ovarian cancer, and other cancers
 CC of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract,
 CC liver, lung, or urogenital. Immune disorders such as Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections may also be treated using, the
 CC proteins and polynucleotides of the invention. Sequences AAC95512 -
 CC AAC95520 and AAB52011 are used in the isolation and characterisation of
 CC the proteins and polynucleotides of the invention.

XX Sequence 385 BP; 121 A; 108 C; 52 G; 103 T; 1 other;

Query Match 2.0%; Score 26; DB 21; Length 385;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
DB 339 TAAAAAAAAAAAAAAAAAAGCTCGAG 364

RESULT 32

AAT84928
ID AAT84928 standard; cDNA; 403 BP.

XX AAT84928;

XX 25-MAR-2003 (updated)
DT 27-APR-1998 (first entry)

XX Human prostate protein HPA16 3' cDNA.

XX Prostate cancer; immunotherapy; therapy; immunodiagnosis; diagnosis;
KM vaccine; human; HPA16; ss.

XX Homo sapiens.

XX MO9733909-A2.

XX 18-SEP-1997.

XX 14-MAR-1997; 97WO-US04192.

XX 15-MAR-1996; 96US-0616745.

XX 10-APR-1996; 96US-0633840.

XX (CORI-) CORIXA CORP.

XX Dillon DC, Reed SG, Twardzik DR;

XX WPI; 1997-470816/43.

XX Immunogenic portions of prostate proteins - useful to develop
PT products to detect, monitor, treat or inhibit development of
PT prostate cancer

XX Claim 28; Page 46; 84pp; English.

XX This sequence comprises a 3' portion of HPA16 cDNA; another
CC isolated clone (see AAT84927) of HPA16 cDNA codes for human prostate
CC protein HPA16 (see AAT23312), an immunogenic portion of which can be
CC used in a claimed pharmaceutical composition for the treatment of
CC prostate cancer. In a vaccine for treatment of prostate cancer, or
CC to raise antibodies suitable for use in diagnosis or monitoring the
CC progression of prostate cancer. HPA16 cDNA was isolated from a
CC human prostate adenocarcinoma cell line Lncap.fgc (ATCC CRL-1740)
CC cell cDNA library by expression screening with human prostatitis
CC sera. DNA sequences (see AAT84927-52) for 17 HPA proteins (see
CC AAT23312-23 and AAT27303-07) are provided and can be used to produce
CC recombinant HPA polypeptides in host cells (particularly E. coli,
CC yeast and mammalian cell lines) and to design primers or probes for
CC use in claimed methods of detecting prostate cancer.
CC (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 403 BP; 135 A; 87 C; 81 G; 100 T; 0 other;

XX Query Match 2.0%; Score 26; DB 18; Length 403;

XX Best Local Similarity 100.0%; Pred. No. 5.5;

XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
DB 378 TAAAAAAAAAAAAAAAAAAGCTCGAG 403

RESULT 33

AA35858
ID AA35858 standard; cDNA; 403 BP.

XX AA35858;

XX 13-JUL-1999 (first entry)

XX cDNA encoding a prostate tumour cell polypeptide.

XX N-terminal; prostate tumour cell; immunogenic; treatment; diagnosis;

XX prostate cancer; ss.

XX Homo sapiens.

XX MO9918210-A2.

XX 15-APR-1999.

XX 07-OCT-1998; 98WO-US21166.

XX 23-JUN-1998; 98US-0102679.

XX 07-OCT-1997; 97US-0946026.

XX (CORI-) CORIXA CORP.

XX Dillon DC, Mitcham JL, Reed SG, Twardzik DR;

XX WPI; 1999-277272/23.

XX New isolated prostate polypeptides useful for the treatment,
PT diagnosis and monitoring of prostate cancer

XX Claim 28; Page 54; 106pp; English.

XX The present sequence encodes a polypeptide obtained from prostate
CC tumour cells. The polypeptide comprises an immunogenic portion of
CC a prostate protein. The polypeptides and DNA obtained from prostate
CC tumour cells, as well as antibodies raised against the protein, can
CC be used in the treatment, diagnosis and monitoring of prostate
CC cancer.

XX Sequence 403 BP; 135 A; 87 C; 81 G; 100 T; 0 other;

XX Query Match 2.0%; Score 26; DB 20; Length 403;

XX Best Local Similarity 100.0%; Pred. No. 5.5;

XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
DB 378 TAAAAAAAAAAAAAAAAAAGCTCGAG 403

RESULT 34

AAC57303
ID AAC57303 standard; DNA; 404 BP.

XX AAC57303;

XX 25-JAN-2001 (first entry)

XX Pinus radiata transcription factor DNA sequence #664.

XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;

XX poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;

XX basic helix-loop-helix zipper; homeotic; homeodomain; MAD5;

XX homeodomain zipper; LIM domain; AP2; ERBBs; zinc finger domain;

XX type 2 CysHis2; CCAAT box element; MYB; ss.
XX Pinus radiata.
XX OS
XX PN W0200053724-A2.

XX 14-SEP-2000.
PD 09-MAR-2000; 2000WO-US06112.
XX 11-MAR-1999; 99US-0266513.
PR 18-AUG-1999; 99US-0149485.
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX Wood M, McGrath A, Shenk MA, Glenn M;
PI WPI; 2000-579369/54.
XX WPI; 2000-579369/54.
XX New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
XX
XX Claim 1; Page 641; 747pp; English.
XX The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC sequence for one such transcription factor. The transcription factor may
CC be used to produce a plant having modified gene expression such as a
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC mahogany species or to modify the activity of a polypeptide in a plant.
CC The transcription factors of the present invention are members from the
CC following families of regulatory proteins: bZIP, bZIP family of G-box
CC binding factors, basic helix-loop-helix zipper,
CC homeoic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements
CC and MYB.
XX
XX Sequence 404 BP; 136 A; 94 C; 100 G; 74 T; 0 other;
SQ
Query Match 2.0%; Score 26; DB 21; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 358 TAAAAAAAAAAAAAAAACTCGAG 383
RESULT 35
AAK8858/c
ID AAK8858 standard; cDNA; 519 BP.
XX
AC AAK8858;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen coding sequence SEQ ID NO: 1174.
XX
KW Human, digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ss.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220563.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 18-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234927.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251956.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WBI, 2001-502630/55.
 DR P-PSDB; AAM93085.
 XX
 XX Polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognosing disorders of the
 PT digestive system, particularly cancer and cancer metastases -
 XX
 PS Claim 1, SEQ ID NO 1174; 986bp; English.
 XX
 CC The present invention provides the protein and coding sequences of a

CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a cDNA encoding a digestive
 CC system antigen of the invention.
 XX
 SQ Sequence 519 BP; 173 A; 86 C; 82 G; 173 T; 5 other;
 Query Match 2.0%; Score 26; DB 22; Length 519;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
 DB 44 TAAAAAAAAAAAAAAAACTCGAG 19
 RESULT 36
 AAS31804/c
 ID AAS31804 standard; cDNA; 519 BP.
 XX
 AC AAS31804;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Human liver associated cDNA polynucleotide #124.
 XX
 KW Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;
 KW chicken; sheep; immunosuppressive; antidiabetic; vasotropic;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
 KW cerebroprotective; neotropic; antibacterial; virucide; fungicide; cancer;
 KW ophthalmological; vulnereary; gene therapy; autoimmune disease; neoplasm;
 KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
 KW cerebrovascular disorder; nervous system disorder; bacterial infection;
 KW fungal infection; viral infection; ocular disorder; endocrine disorder;
 KW gastrointestinal disorder; renal disorder; respiratory disorder;
 KW wound healing; skin aging; organ transplantation; tissue regeneration;
 anti-fertility.
 XX
 OS Homo sapiens.
 XX
 PN WO200155355-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01351.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246569.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-457728/49.
XX P-PDB; AAU20093.
XX
XX Isolated nucleic acid molecule encoding a human liver related protein
PT is used in preventing, treating or ameliorating disorders of the liver
PT particularly cancer of the liver -
XX
XX
PS Claim 1; SEQ ID No 134; 526pp; English.
XX
XX Sequences AA531681-AA531926 represent cDNA molecules, which encode the
CC liver associated polypeptides of the invention. Liver associated
CC polypeptides and their associated polynucleotides are useful in the
CC diagnosis, treatment and prevention of various types of disorders in e.g.
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
CC pathological condition can be determined by detecting the presence or
CC absence of a mutation in a liver associated polynucleotide. The treatable
CC disorders include autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders such as the breast or liver,
CC cardiovascular disorders such as cardiac arrest, cerebrovascular
CC disorders such as cerebral ischaemia, nervous system disorders such as
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma and pleurisy. The polypeptides can
CC also be used to aid wound healing, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, to regenerate tissues and in
CC chemotaxis.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO

CC at fcp.wipo.int/pub/published_pct_sequences.
Query Match 2.0%; Score 26; DB 22; Length 519;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAAAATCTCGAG 1323
Db 44 TAAAAAAAAAAAAAAAAAATCTCGAG 19
RESULT 37
ABN90159/c
ID ABN90159 standard; cDNA; 519 BP.
XX
AC ABN90159;
XX
XX 24-JUL-2002 (first entry)
DE Human liver antigen HPAIC84 cDNA, SEQ ID NO:134.
XX
XX Human liver antigen; liver disorder; hepatic disorder; infection;
KM hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;
KM cirrhosis; granulomatous hepatitis; toxin damage; drug damage;
KM autoimmune disease; Wilson's disease; primary biliary cirrhosis;
KM neoplastic disorder; cancer; tumor; portal hypertension;
KM gastrointestinal disorder; hepatitis; drug screening; gene therapy;
KM chromosome mapping; forensic analysis; antibody preparation;
KM hepatocellular carcinoma; cytotoxic; antiinflammatory; virulence; antibacterial;
KM fungicide; parasiticide; antidote; immunosuppressive; gene; ss.
XX
OS Homo sapiens.
XX
XX US2002042096-A1.
XX
XX 11-APR-2002.
XX
PF 17-JAN-2001; 2001US-0764887.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229510P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 27-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.

PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239355P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC,
XX
XX WPI; 2002-381944/41.
DR P-PSDB; ABP40954.
XX
XX New nucleic acid encoding human liver antigens, useful for diagnosis,
PT treatment and prevention of e.g. hepatitis and hepatic cancer, also
PT related polypeptides and antibodies -
XX
XX Claim 1; SEQ ID No 134; 181bp; English.
XX
XX The invention relates to 145 novel human liver antigens (ABP40831-
CC ABP40975) and to cDNAs encoding them (ABN90036-ABN90180), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human liver antigen
CC polynucleotides, antibodies against human liver antigens, and the use of
CC liver antigen polynucleotides and polypeptides in diagnosing, treating,
CC prognosing or preventing various disorders of the liver. Such conditions
CC include viral infections (e.g., cytomegalovirus, Epstein-Barr virus,
CC hepatitis A virus, hepatitis B virus and hepatitis C virus), parasitic
CC infections (e.g., Clonorchis sinensis, Echinococcus granulosus and
CC Entamoeba histolytica), and also bacterial and fungal infections. Other
CC disorders that may be treated include inflammatory conditions (e.g.,
CC cirrhosis and granulomatous hepatitis), damage caused by drugs or toxins,
CC autoimmune diseases (e.g., Wilson's disease, primary biliary cirrhosis),
CC neoplastic disorders (e.g., adenomas, haemangiomas and hepatocellular
CC carcinoma), portal hypertension, or gastrointestinal disorders (e.g.,
CC peptic ulcers, gastritis and peritoneal diseases). Liver antigen
CC polypeptides and polynucleotides may also be used in screening for
CC compounds which modulate liver antigen expression or activity. The
CC polynucleotides may further be used for gene therapy, chromosome
CC mapping, in the identification of individuals and in forensic analysis,
CC and the polypeptides may be used as molecular weight markers or to
CC prepare antibodies useful in disease diagnosis, drug targeting and
CC phenotyping. The present sequence represents cDNA encoding a human liver
CC antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO at seqdata.uspto.gov/sequence/
XX

SO Sequence 519 BP; 173 A; 86 C; 82 G; 173 T; 5 other;
Query Match 2.0%; Score 26; DB 24; Length 519;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAAAATCTCGAG 1323
Db 44 TAAAAAAAAAAAAAAAAAATCTCGAG 19

RESULT 38
AAS34097/c
ID AAS34097 standard; CDNA: 524 BP.
XX
AC AAS34097;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human cDNA encoding a novel foetal antigen, SEQ ID No 621.
XX
KW Human; foetal tissue antigen; ss; antiinflammatory; neuroprotective;
KW immunomodulator; cardiovascular; cytosolic; nephrotoxic;
KW cardiovascular; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; breast neoplasm; cancer;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; food additive.
XX
OS Homo sapiens.
XX
PN WO200155312-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01321.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 05-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 02-OCT-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.


```

XX RESULT 40
XX ID AAH87896 standard; cDNA, 577 BP.
XX AC AAH87896;
XX DT 25-SEP-2001 (first entry)
XX DE Peppermint plant oil gland expressed cDNA 252.
XX KW Peppermint; plant oil gland cell; terpenoid essential oil; resin;
XX genetic mapping; antisense suppression; recombinant expression; ss.
XX OS Mentha x piperita.
XX WO200153319-A1.
XX PD 26-JUL-2001.
XX PF 19-JAN-2001; 2001WO-US02567.
XX PR 20-JAN-2000; 2000US-0177264.
XX PA (CROT/) CROTEAU R B.
XX PA (LANG/) LANGE B M.
XX PA (WILD/) WILDUNG M R.
XX PI Croteau RE, Lange BM, Wildung MR;
XX WP1; 2001-488706/53.
XX PT New nucleic acid molecules corresponding to mRNA molecules expressed in
XX PT peppermint oil glands for enhancing expression of plant oil gland cell
XX PT proteins -
XX PS Claim 1; Page 172; 251pp; English.
XX CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that
XX CC correspond to all or part of a mRNA molecule expressed in plant oil
XX CC gland cells, especially peppermint and plant oil glands that produce
XX CC terpenoid essential oils and resins. The nucleic acids are useful for
XX CC genetically mapping a plant genome for genes expressed in plant oil
XX CC gland cells and to suppress (for example by antisense suppression) or
XX CC enhance their expression (for example by genetically transforming a
XX CC plant cell with a replicable expression vector that expresses one or more
XX CC proteins naturally expressed in plant oil gland cells). The nucleic acids
XX CC are also useful for recombinant expression of plant oil gland proteins
XX CC required for terpenoid essential oil and/or resin production in bacterial
XX CC and/or yeast cells.
XX SO Sequence 577 BP; 185 A; 117 C; 130 G; 145 T; 0 other;

Query Match 2.0%; Score 26; DB 22; Length 577;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 519 TAAAAAAAAAAAAAAAAACTCGAG 544

RESULT 41
XX ID AAH34095 standard; cDNA, 591 BP.
XX AC AAH34095;
XX DT 03-SEP-2001 (first entry)
XX DE Human colon cancer antigen encoding cDNA SEQ ID NO:1177.
XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

```

KM		colorectal carcinoma; ss.
XX		
OS	Homo sapiens.	
XX		
PN	WO200122920-A2.	
XX		
PD	05-APR-2001.	
XX		
PF	28-SEP-2000; 2000OWO-US26524.	
XX		
PR	29-SEP-1999; 99US-0157137.	
PP	03-NOV-1999; 99US-0163280.	
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
F1	Ruben SM, Barash SC, Birse CE, Rosen CA;	
XX		
DR	WPI; 2001-235357/24.	
DR	P-PSDB; AAG74690.	
XX		
PT	Nucleic acid encoding 4277 human colon cancer-associated polypeptides,	
FT	useful for preventing, diagnosing and/or treating colorectal cancers -	
PS		
XX	Claim 1; Page 2977-2978; 9803bp; English.	
CC	AAB32943 to AAH37195 and AAG73514 to AAG77788 represent human colon	
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where	
CC	the proteins are collectively known as colon cancer antigens. The colon	
CC	cancer antigens have cytostatic activity and can be used in gene	
CC	therapy and vaccine production. N and P may be used in the prevention,	
CC	diagnosis and treatment of diseases associated with inappropriate P	
CC	expression. For example, N and P may be used to treat disorders	
CC	associated with decreased expression by rectifying mutations or deletions	
CC	in a patient's genome that affect the activity of P by expressing P.	
CC	Inactive proteins or to supplement the patients own production of P.	
CC	Additionally, N may be used to produce the colon cancer-associated Ps,	
CC	by inserting the nucleic acids into a host cell and culturing the cell	
CC	to express the proteins. N and P can be used in the prevention, diagnosis	
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204	
CC	and AAB77789 represent sequences used in the exemplification of the	
CC	present invention.	
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were	
CC	missing at time of publication, meaning no sequences are present for	
CC	SEQ ID NO:1027 to 1052, 7921 and 7922.	
SQ		
XX	Sequence 591 BP; 199 A; 95 C; 114 G; 182 T; 1 other;	
Query Match	2.0%; Score 26; DB 22; Length 591;	
Best Local Similarity	100.0%; Pred. No. 5.2;	
Matches	26; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	1298 TAAAAAAAAAAAAAACTCGAG 1323	
Db	545 TAAAAAAAAAAAAAACTCAG 570	
RESULT 42		
ID	AAH8017/C	
AC	AAH8017 standard; cDNA; 639 BP.	
XX		
DT	25-SEP-2001 (first entry)	
XX		
DE	Peppermint plant oil gland expressed cDNA 373.	
XX		
KM	Peppermint; plant oil gland cell; terpenoid essential oil; resin;	
KM	genetic mapping; antisense suppression; recombinant expression; ss.	
XX		
OS	Mentha x piperita.	
XX		
PN	WO200153319-A1.	
XX		

PD 26-JUL-2001.
XX
XX 19-JAN-2001; 2001MO-US802567.
XX
XX 20-JAN-2000; 2000US-0177264.
XX
XX (CROT/) CROTEAU R B.
XX (LANG/) LANGE B M.
XX (WILD/) WILDUNG M R.
XX
XX Croteau RB, Lange BM, Wildung MR;
XX MPI; 2001-488706/53.
XX
XX New nucleic acid molecules corresponding to mRNA molecules expressed in
PT peppermint oil glands for enhancing expression of plant oil gland cell
PT proteins -
XX
XX Claim 1; Page 214; 251pp; English.
XX
XX The invention relates to nucleic acid molecules (AAH97645-AAH88116) that
CC correspond to all or part of a mRNA molecule expressed in plant oil
CC gland cells, especially peppermint and plant oil glands that produce
CC terpenoid essential oils and resins. The nucleic acids are useful for
CC genetically mapping a plant genome for genes expressed in plant oil
CC gland cells and to suppress (for example by antisense suppression) or
CC enhance their expression (for example by genetically transforming a
CC plant cell with a replicable expression vector that expresses one or more
CC proteins naturally expressed in plant oil gland cells). The nucleic acids
CC are also useful for recombinant expression of plant oil gland proteins
CC required for terpenoid essential oil and/or resin production in bacterial
CC and/or yeast cells.
XX
XX Sequence 639 BP; 191 A; 118 C; 137 G; 193 T; 0 other;
SO
Query Match 2.0%; Score 26; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
DB 31 TAAAAAAAAAAAAAAAAAATCGAG 6
RESULT 43
AAH33451
ID AAH33451 standard; cDNA; 646 BP.
XX
XX AAH33451;
XX
XX 03-SEP-2001 (first entry)
XX
XX Human colon cancer antigen encoding cDNA SEQ ID NO:507.
XX
XX Human colon cancer antigen encoding cDNA; diagnosis; detection;
XX colorectal carcinoma; ss.
XX
XX Homo sapiens.
XX
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000MO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
XX 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX MPI; 2001-235357/24.
XX
XX

DR P-PSDB; AAG74020.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1; Page 2586; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG7789 represent sequences used in the exemplification of the
CC present invention. 682 and page 7053 of the sequence listing were
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 646 BP; 144 A; 158 C; 179 G; 161 T; 4 other;
SO
Query Match 2.0%; Score 26; DB 22; Length 646;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
DB 569 TAAAAAAAAAAAAAAAAAATCGAG 594
RESULT 44
AAC93368
ID AAC93368 standard; cDNA; 655 BP.
XX
XX AAC93368;
XX
XX 26-FEB-2001 (first entry)
XX
XX Human secreted protein gene 5 SEQ ID NO:15.
XX
XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
XX antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
XX cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
XX fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
XX hyperproliferative disorder; cardiovascular disorder; angiogenesis;
XX cerebrovascular disorder; nervous system disorder; ocular disorder;
XX wound healing; skin aging; food additive; preservative; ss.
XX
XX Homo sapiens.
XX
XX WO200061620-A1.
XX
XX 19-OCT-2000.
XX
XX 06-APR-2000; 2000MO-US09069.
XX
XX 09-APR-1999; 99US-0128702.
XX 20-JAN-2000; 2000US-0177049.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX MPI; 2000-619225/59.
XX
XX

DR P-PSDB; AAB51624.
XX Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition -
XX
XX Claim 1; Page 424-425; 540pp; English.
XX
XX The polynucleotide sequences given in AAC93364 to AAC93412 encode the
CC human secreted proteins given in AAB51620 to AAB51668. AAB51669 to
CC AAB51722 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC antiarthritic; immunosuppressive; antineumatic; antiproliferative;
CC cytosratic; cardiant; vasotropic; cerebroprotective; nootropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC optalmimological. The polynucleotides and polypeptides can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders, cardiovascular
CC disorders, cerebrovascular disorders, angiogenesis, nervous system
CC disorders, infections caused by bacteria, viruses and fungi and ocular
CC disorders. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. AAG93355 to AAC93363 and
CC AAB51619 represent sequences which are used in the exemplification of the
CC present invention.
XX
SQ Sequence 655 BP; 192 A; 133 C; 135 G; 190 T; 5 other;

Query Match 2.0%; Score 26; DB 21; Length 655;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
Db 627 TAAAAAAAAAAAAAAAAAAGCTCGAG 652

RESULT 45
ABZ67172 ID ABZ67172 standard; cDNA; 655 BP.
XX
XX ABZ67172;
AC
XX
XX
DT 26-MAR-2003 (first entry)
XX
DE Human secreted protein encoding cDNA SEQ ID NO 292.
XX
XX Human; secreted protein; nootropic; neuroprotective; cytosratic;
XX Humanc; dematologic; immunosuppressive; antiinflammatory; anti-HIV;
XX vlnetary; antibacterial; antiparkinsonian; antisticking; antianemic;
XX antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
XX antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
XX antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
XX cardiovascular disorder; neurological disease; nephrotropic;
XX gene therapy; gene; ds.
OS Homo sapiens.
XX
XX
XX WO200277186-A2.
PN
XX
XX 03-OCT-2002.
PD
XX
XX 26-MAR-2002; 2002WO-US09188.
PF
XX
XX 27-MAR-2001; 2001US-278650P.
PR 12-SEP-2001; 2001US-0950082.

PR 12-SEP-2001; 2001US-0950083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Ruben SM;
PI
XX
XX WPI; 2003-040583/03.
DR
DR P-PSDB; ABP99751.
XX
XX
XX New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGAC19), useful for preventing, treating or diagnosing e.g.
PT AIDS, multiple sclerosis, herpes virus, leukemia, tick-borne
PT encephalitis or West Nile fever -
PS
PS Claim 7; Page 1321; 2423pp; English.
XX
XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections.
XX
SQ Sequence 655 BP; 192 A; 133 C; 135 G; 190 T; 5 other;

Query Match 2.0%; Score 26; DB 25; Length 655;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
Db 627 TAAAAAAAAAAAAAAAAAAGCTCGAG 652

RESULT 46
AAD08196 ID AAD08196 standard; cDNA; 682 BP.
XX
XX AAD08196;
AC
XX
XX
DT 07-AUG-2001 (first entry)
XX
DE Human secreted protein-encoding gene 6 cDNA clone HE90W91, SEQ ID NO.16.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiogenic disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder; gene therapy;
XX endocrine disorder; infection; wound healing; vlnetary;
XX cell culture; chemotaxis; food additive;
XX binding partner identification; ss.
OS Homo sapiens.
XX
XX
XX
XX Key Location/Qualifiers
FH CDS 33..509
FT /*tag= a
FT /product= "Human secreted protein precursor"
FT 33..113
FT /*tag= b

mat_peptide 114..506
 /tag= C
 /product= "Mature human secreted protein precursor"
 MO200132837-A1.
 10-MAY-2001.
 17-OCT-2000; 2000MO-US28664.
 02-NOV-1999; 99US-0163085.
 17-DEC-1999; 99US-0172411.
 (HUMA-) HUMAN GENOME SCI INC.
 Piscaglia M, Mel P, Lafleur DW, Olsen HS, Baker K, Ebner R;
 Komatsoulis G, Rosen CA, Ruben SM, Duan RD, Young PE, Florence KA;
 Moore PA, Birse CE, Ni J, Soppet DR, Shi Y;
 MPI; 2001-328782/34.
 P-PSDB; AAEO3769.
 Novel human secreted proteins and nucleic acids for diagnosing,
 preventing and treating neurological, cardiovascular, infectious,
 autoimmune, gastrointestinal, bone disorders, cancer, particularly
 ovarian cancer -
 Claim 1; Page 378; 421pp; English.
 AAD08191-AAD08213 represent cDNAs corresponding to 19 human secreted
 protein genes and AAEO3764-AAEO3786 represent the proteins they encode.
 AAEO3787-AAEO3800 represent human secreted protein fragments or variants.
 The genes and their secreted proteins are useful for preventing,
 treating or ameliorating medical conditions, e.g., by protein or gene
 therapy. Pathological conditions can be diagnosed by determining the
 amount of the new protein in a sample or by determining the presence of
 mutations in the new genes. Specific uses are described for each of the
 19 genes, based on the tissues in which they are most highly expressed,
 and include developing products for the diagnosis or treatment of
 proliferative disorders, cancer, tumours, foetal and developmental
 abnormalities, haematopoietic disorders, diseases of the immune system,
 AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 allergies, neurological disorders (e.g., Alzheimer's disease,
 Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 cardiovascular disorders, angiotensin disorders, kidney disorders,
 gastrointestinal disorders, pregnancy-related disorders, endocrine
 disorders, and infections. The proteins can also be used to aid wound
 healing and epithelial cell proliferation, to prevent skin aging due to
 sunburn, to maintain organs before transplantation, for supporting cell
 culture of primary tissues, to regenerate tissues, to identify their
 cognate ligands or binding partners, and in chemotaxis, and can be used
 as a food additive or preservative to modify storage properties.
 Antibodies specific for a protein of the invention can be used in
 alleviating symptoms associated with the disorders mentioned above, and
 in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 immunosorbent assay (ELISA). The present sequence represents a human
 secreted protein-encoding cDNA of the invention.
 Sequence 682 BP; 149 A; 172 C; 222 G; 139 T; 0 other;
 Query Match 2.0%; Score 26; DB 22; Length 682;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1298 TAAAAAATCTGAG 1323
 Db 644 TAAAAAATCTGAG 669
 RESULT 47
 AA298306
 ID AA298306 standard; DNA; 695 BP.

AA298306;
 14-JUN-2000 (first entry)
 A. thaliana HSP 17.6A protein encoding DNA.
 Environmental stress; plant; transgenic plant; anaerobic; flooding; cold;
 dehydration; drought; heat stress; salinity; osmotolerance; ds.
 Arabidopsis thaliana.
 WO200008187-A2.
 17-FEB-2000.
 04-AUG-1999; 99WO-EP05652.
 04-AUG-1998; 98EP-0202634.
 (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
 Lee JH, Verbruggen N;
 MPI; 2000-205726/18.
 P-PSDB; AAY77926.
 Isolation of polynucleic acids useful for producing transgenic plant by
 isolating genes involved in tolerance to environmental stress -
 Claim 8; Page 84; 312pp; English.
 The invention relates to isolation of coding sequences and/or genes
 involved in tolerance to environmental stress in plants. The sequences
 (AA298305-298365) are useful for producing a transgenic plant having
 enhanced tolerance or resistance to environmental stress conditions such
 as anaerobic, flooding, cold, dehydration, drought, heat stress or
 salinity. This is useful for producing improved yield, growth,
 development and productivity under environmental stress conditions, and
 also provides growth of crops in areas where they cannot grow without
 the induced osmotolerance. Sequences AA298305-365 represent
 CC polynucleotide sequences from A. thaliana that are involved in
 CC environmental stress tolerance.
 Sequence 695 BP; 221 A; 124 C; 178 G; 172 T; 0 other;
 Query Match 2.0%; Score 26; DB 21; Length 695;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1298 TAAAAAATCTGAG 1323
 Db 669 TAAAAAATCTGAG 694
 RESULT 48
 AA217199/C
 ID AA217199 standard; cDNA; 712 BP.
 AA217199;
 12-OCT-1999 (first entry)
 Human gene expression product cDNA sequence SEQ ID NO:4670.
 Human; gene; gene expression product; diagnosis; therapy; probe;
 detection; mapping; tissue typing; profiling; forensic; cancer;
 genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 Homo sapiens.
 MO938972-A2.
 XX

PD 05-AUG-1999.
 XX
 PF 28-JAN-1999; 99WO-US01619.
 XX
 PR 03-APR-1998; 98US-0080666.
 PR 28-JAN-1998; 98US-0072910.
 PR 24-FEB-1998; 98US-0075954.
 PR 31-MAR-1998; 98US-0080114.
 PR 03-APR-1998; 98US-0080515.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HSE-) HYSEQ INC.
 XX
 PI Ctkvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo U, Garcia PD, Garcia V, Giese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Grain B, Sudduth-Klinger J, Williams LT;
 XX
 DR WPI; 1999-494092/41.
 XX
 PT Novel human genes and their expression products which are
 PT differentially expressed in different cell types
 XX
 PS Claim 1; Page 2218; 2479pp; English.
 XX
 CC The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AA212532 to AA217779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The
 CC polynucleotides can be used as a source of primers and probes, which can
 CC be used for a variety of purpose, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensics, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and
 CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels between
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
 CC identify a genetic predisposition or susceptibility to a disease such as
 CC cancer). The polynucleotides of the invention are especially used in the
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
 CC and lung cancer. The polynucleotides can also be used to screen for
 CC peptide analogues and antagonists.
 CC
 XX
 SQ Sequence 712 BP; 250 A; 130 C; 96 G; 236 T; 0 other;
 XX
 Query Match 2.0%; Score 26; DB 20; Length 712;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1298 TAAAAAAAAAAAAAAAAAATCTCGAG 1323
 Db 91 TAAAAAAAAAAAAAAAAAATCTCGAG 66
 XX
 RESULT 49
 AAC63453
 ID AAC63453 standard; cDNA; 735 BP.
 XX
 AC AAC63453;
 XX
 DT 09-FEB-2001 (first entry)
 XX
 DE Human secreted protein coding sequence SEQ ID NO: 54.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; human; secreted protein; ss.
 XX
 OS Homo sapiens.
 XX
 FN W0200061779-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 06-APR-2000; 2000WO-US09068.
 XX
 PR 09-APR-1999; 99US-0128699.
 PR 20-JAN-2000; 2000US-0177050.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX
 DR WPI; 2000-647424/62.
 DR P-PSDB; AAB29845.
 XX
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 XX
 PS Claim 1; Page 422; 495pp; English.
 XX
 CC The invention relates to the isolation of genes AAC63410-C63458 encoding
 CC 49 human secreted proteins AAB29802-B29850. The genes can be used to
 CC generate fusion proteins by linking to the gene for the human
 CC immunoglobulin G Fc portion (SEQID1) for increasing the stability of
 CC the fusion protein as compared to the human protein only. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated
 CC from a range of human tissues disclosed in the specification. The
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischemias; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC
 XX
 SQ Sequence 735 BP; 170 A; 150 C; 143 G; 265 T; 7 other;
 XX
 Query Match 2.0%; Score 26; DB 21; Length 735;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1298 TAAAAAAAAAAAAAAAAAATCTCGAG 1323
 Db 657 TAAAAAAAAAAAAAAAAAATCTCGAG 682
 XX
 RESULT 50
 AAC66438
 ID AAC66438 standard; DNA; 784 BP.
 XX
 AC AAC66438;
 XX
 DT 14-FEB-2001 (first entry)
 XX
 DE Human secreted protein coding sequence SEQ ID NO: 39.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; ss.
 XX
 OS Homo sapiens.

XX WO200058350-A1.
PN
XX
XX 05-OCT-2000.
PD
XX
XX 22-MAR-2000; 2000WO-US07483.
PF
XX
XX 26-MAR-1999; 99US-0126596.
PR
XX 22-DEC-1999; 99US-0171552.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsu G;
PI
XX
XX WPI; 2000-602357/57.
DR
XX P-PSDB; AAB32030.
PT
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX
XX
PS Claim 1; Page 350; 423pp; English.
XX
XX The invention relates to the isolation of genes AAC66410-C66458 encoding
CC 49 human secreted proteins AAB32002-B32050. The genes can be used to
CC generate fusion proteins by linking to the gene for the human
CC immunoglobulin G Fc portion (SEQID) for increasing the stability of
CC the fusion protein as compared to the human protein only. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated
CC from a range of human tissues disclosed in the specification. The
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX
XX Sequence 784 BP; 220 A; 169 C; 139 G; 251 T; 5 other;
SQ

Query March 2.0%; Score 26; DB 21; Length 784;
Best Local Similarity 100.0%; Freq. No. 4.9;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
|||
DB 759 TAAAAAAAAAAAAAAAACTCGAG 784

Search completed: January 28, 2004, 18:18:55
Job time : 406 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 15:56:05 ; Search time 109 Seconds
(without alignments)
5357.338 Million cell updates/sec

Title: US-09-881-556A-3
Perfect score: 1323
Sequence: 1 gaattcgagcgagattctt.....aaaaaaaaaaactcgag 1323

Scoring table: OLIGO_NIC
Gapop 60.0, Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 15

Total number of hits satisfying chosen parameters: 15432

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Issued Patente_NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result					Description	
No.	Score	Query Match	Length	ID		
1	2.2	1049	4	US-09-149-476-58	Sequence 58, App1	
2	2.0	1017	3	US-08-849-751-1	Sequence 1, App1	
3	2.0	1017	4	US-09-478-816-1	Sequence 1, App1	
4	2.0	1061	4	US-09-370-838-149	Sequence 149, App1	
5	2.7	2.0	1390	3	US-08-605-150A-7	Sequence 7, App1
6	2.0	317	4	US-09-370-838-17	Sequence 17, App1	
7	2.0	403	3	US-08-946-026-10	Sequence 10, App1	
8	2.0	815	4	US-09-439-554-21	Sequence 21, App1	
9	2.0	1126	4	US-09-489-847-60	Sequence 60, App1	
10	2.0	1141	4	US-09-674-5298-7	Sequence 7, App1	
11	2.0	1153	4	US-08-747-562-11	Sequence 11, App1	
12	2.0	1131	4	US-09-370-838-27	Sequence 27, App1	
13	2.0	1352	2	US-08-415-593-44	Sequence 44, App1	
14	2.0	1354	4	US-09-149-476-207	Sequence 207, App1	
15	2.0	1486	4	US-09-149-476-23	Sequence 23, App1	
16	2.0	1631	4	US-09-149-476-97	Sequence 97, App1	
17	2.0	2379	4	US-09-205-358-175	Sequence 175, App1	
18	2.0	3066	3	US-09-086-912-1	Sequence 1, App1	
19	2.0	3066	4	US-09-203-453-1	Sequence 1, App1	
20	2.0	3066	4	US-09-900-236-1	Sequence 1, App1	
21	2.0	3767	3	US-08-928-941D-28	Sequence 28, App1	
22	2.0	3767	3	US-08-928-941D-30	Sequence 30, App1	
23	2.0	3767	4	US-09-280-590A-28	Sequence 28, App1	
24	2.0	3767	4	US-09-280-590A-30	Sequence 30, App1	
25	2.0	4382	4	US-09-023-905A-3	Sequence 3, App1	
26	1.9	38	3	US-09-454-704A-2	Sequence 2, App1	
27	1.9	38	3	US-09-454-704A-12	Sequence 12, App1	

28	25	1.9	42	3	US-09-454-704A-14	Sequence 14, App1
29	25	1.9	247	1	US-08-594-031-103	Sequence 103, App
30	25	1.9	289	1	US-08-341-568-3	Sequence 3, App1
31	25	1.9	289	2	US-08-911-020-3	Sequence 3, App1
32	25	1.9	307	4	US-09-091-725-28	Sequence 28, App1
33	25	1.9	319	1	US-08-341-568-5	Sequence 5, App1
34	25	1.9	319	2	US-08-911-020-5	Sequence 5, App1
35	25	1.9	333	3	US-08-946-026-35	Sequence 35, App1
36	25	1.9	340	3	US-08-888-077A-39	Sequence 39, App1
37	25	1.9	360	3	US-08-888-077A-37	Sequence 37, App1
38	25	1.9	375	3	US-08-946-026-23	Sequence 23, App1
39	25	1.9	376	4	US-09-220-132-128	Sequence 128, App
40	25	1.9	377	3	US-08-946-026-41	Sequence 41, App1
41	25	1.9	395	4	US-09-370-838-164	Sequence 164, App
42	25	1.9	490	2	US-08-630-822A-92	Sequence 92, App1
43	25	1.9	490	2	US-09-005-069-92	Sequence 92, App1
44	25	1.9	490	4	US-09-171-156A-41	Sequence 41, App1
45	25	1.9	490	4	US-09-004-730A-41	Sequence 41, App1
46	25	1.9	490	4	US-08-981-799A-41	Sequence 41, App1
47	25	1.9	506	4	US-09-370-838-263	Sequence 263, App
48	25	1.9	536	1	US-08-341-568-1	Sequence 1, App1
49	25	1.9	536	2	US-08-911-020-1	Sequence 1, App1
50	25	1.9	556	4	US-09-370-838-242	Sequence 242, App
51	25	1.9	575	4	US-09-205-258-219	Sequence 219, App
52	25	1.9	616	2	US-08-630-822A-93	Sequence 93, App1
53	25	1.9	616	2	US-09-005-069-93	Sequence 93, App1
54	25	1.9	616	4	US-09-171-156A-42	Sequence 42, App1
55	25	1.9	616	4	US-09-004-730A-42	Sequence 42, App1
56	25	1.9	616	4	US-08-981-799A-42	Sequence 42, App1
57	25	1.9	634	4	US-09-669-751-2	Sequence 2, App1
58	25	1.9	642	4	US-09-370-838-119	Sequence 119, App
59	25	1.9	680	4	US-09-489-847-73	Sequence 73, App1
60	25	1.9	684	4	US-08-226-264-27	Sequence 27, App1
61	25	1.9	685	2	US-08-254-354-1	Sequence 1, App1
62	25	1.9	685	2	US-08-354-354-3	Sequence 3, App1
63	25	1.9	685	5	PCT-US95-06137-1	Sequence 1, App1
64	25	1.9	685	5	PCT-US95-06137-3	Sequence 3, App1
65	25	1.9	748	4	US-09-372-422A-45	Sequence 45, App1
66	25	1.9	748	4	US-09-004-615-55	Sequence 55, App1
67	25	1.9	794	3	US-09-188-930-41	Sequence 41, App1
68	25	1.9	794	4	US-09-312-283C-41	Sequence 41, App1
69	25	1.9	813	4	US-09-370-838-29	Sequence 29, App1
70	25	1.9	839	1	US-08-652-859-1	Sequence 1, App1
71	25	1.9	839	1	US-08-919-706-1	Sequence 1, App1
72	25	1.9	839	2	US-09-153-751-1	Sequence 1, App1
73	25	1.9	842	4	US-09-149-476-115	Sequence 115, App
74	25	1.9	848	4	US-09-370-838-160	Sequence 160, App
75	25	1.9	855	3	US-09-060-726A-1	Sequence 1, App1
76	25	1.9	856	3	US-09-060-726A-3	Sequence 3, App1
77	25	1.9	860	1	US-08-181-271A-33	Sequence 33, App1
78	25	1.9	860	1	US-08-449-315-33	Sequence 33, App1
79	25	1.9	860	1	US-08-444-803-33	Sequence 33, App1
80	25	1.9	860	1	US-08-449-043-33	Sequence 33, App1
81	25	1.9	860	1	US-08-456-265A-33	Sequence 33, App1
82	25	1.9	860	1	US-08-455-416-33	Sequence 33, App1
83	25	1.9	860	1	US-08-455-244-33	Sequence 33, App1
84	25	1.9	860	1	US-08-454-876-33	Sequence 33, App1
85	25	1.9	860	2	US-08-457-364-33	Sequence 33, App1
86	25	1.9	860	2	US-08-456-262-33	Sequence 33, App1
87	25	1.9	860	2	US-08-456-240-33	Sequence 33, App1
88	25	1.9	860	2	US-08-455-736-33	Sequence 33, App1
89	25	1.9	860	2	US-08-971-217-33	Sequence 33, App1
90	25	1.9	860	2	US-09-350-600-33	Sequence 33, App1
91	25	1.9	867	4	US-09-262-537-13	Sequence 13, App1
92	25	1.9	893	4	US-09-370-838-166	Sequence 166, App
93	25	1.9	897	3	US-08-899-330-14	Sequence 14, App1
94	25	1.9	906	4	US-09-489-847-72	Sequence 72, App1
95	25	1.9	911	4	US-09-461-325-63	Sequence 63, App1
96	25	1.9	917	4	US-09-227-357-123	Sequence 123, App
97	25	1.9	917	2	US-08-951-822-1	Sequence 1, App1
98	25	1.9	917	4	US-09-368-951-1	Sequence 1, App1
99	25	1.9	917	4	US-09-229-947-1	Sequence 1, App1
100	25	1.9	945	4	US-09-482-273-53	Sequence 53, App1

ALIGNMENTS

```
RESULT 1
US-09-149-476-58
; Sequence 58, Application US/09149476
; Patent No. 6420526
GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149, 476
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/056,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
```

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

```

Query Match      2.2%: Score 29; DB 4; Length 1049;
Best Local Similarity 100.0%: Pred. No. 0.01;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Qy      1295 AAGTAAAAAAAAAAAAAAAAAAGCTCGAG 1323
        |||||||||||||||||||||
Db      1012 AAGTAAAAAAAAAAAAAAAAAAGCTCGAG 1040

```

Sequence 1, Application US/08849751
Patent No. 6190890

GENERAL INFORMATION:

APPLICANT: VAN DEN BROECK, HENRIETTE C.
APPLICANT: DE GRAAF, LEEBDEERT H.
APPLICANT: VISSER, JACOB
APPLICANT: VAN COVEN, ALBERT J.J.
TITLE OF INVENTION: FUNGAL CELLULASES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW, suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastISO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,751
FILING DATE: 12-JUN-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/04511
FILING DATE: 14-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 24615-20094.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-822-0168
TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1017 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Aspergillus niger
STRAIN: M400
INDIVIDUAL ISOLATE: CBS120.49

FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 57...773
OTHER INFORMATION: product="Cellulase"

US-08-849-751-1

```

Query Match 2.0%; Score 27; DB 3; Length 1017;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1297 GTAAAAAAAAAAAAAAAAAATCTGAG 1323
        |||
Db       991 GTAAAAAAAAAAAAAAAAAATCTGAG 1017

RESULT 3
US-09-478-816-1
; Sequence 1, Application US/09478816
; Patent No. 6306635
; GENERAL INFORMATION:
; APPLICANT: VAN DEN BROECK, HENRIETTE C.
; APPLICANT: DE GRAEF, LEENDERT H.
; APPLICANT: VISSER, JACOB
; APPLICANT: VAN OOTEN, ALBERT J. J.

```

TITLE OF INVENTION: FUNGAL CELLULASES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW, suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/478,816
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/849,751
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: PCT/EP96/04511
FILING DATE: 14-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 24615-20094.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-822-0168
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1017 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Aspergillus niger
STRAIN: N400
INDIVIDUAL ISOLATE: CBS120.49
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 57..773
OTHER INFORMATION: product="Cellulase"
US-09-478-816-1

Query Match 2.0%; Score 27; DB 4; Length 1017;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
Db 991 GTAAAAAAAAAAAAAAAAAACTCGAG 1017

RESULT 4
US-09-370-838-149
Sequence 149, Application US/09370838
Patent No. 6444425
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamath, Roadoh
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
FILE REFERENCE: 210121.475C1
CURRENT APPLICATION NUMBER: US/09/370,838
CURRENT FILING DATE: 1999-08-09

EARLIER APPLICATION NUMBER: US 09/285,323
EARLIER FILING DATE: 1999-04-02
NUMBER OF SEQ ID NOS: 289
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 149
LENGTH: 1061
TYPE: DNA
ORGANISM: Homo sapien
US-09-370-838-149

Query Match 2.0%; Score 27; DB 4; Length 1061;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
Db 1035 GTAAAAAAAAAAAAAAAAAACTCGAG 1061

RESULT 5
US-08-605-150A-7
Sequence 7, Application US/08605150A
Patent No. 6103520
GENERAL INFORMATION:
APPLICANT: Topfer, Reinhard
APPLICANT: Hausmann, Ludwig
APPLICANT: Schell, Josef
TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klein & Szekeres
STREET: 4199 Campus Drive, Suite 700
CITY: Irvine
STATE: CA
COUNTRY: USA
ZIP: 92715
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,150A
FILING DATE: 01-MAR-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02936
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P4329827.3
FILING DATE: 03-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Szekeres, Gabor L.
REGISTRATION NUMBER: 28,675
REFERENCE/DOCKET NUMBER: 542-04-PA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-854-5502
TELEFAX: 714-854-4897
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1390 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cuphea lanceolata
IMMEDIATE SOURCE:
LIBRARY: ZAP cDNA library
CLONE: CIGPDH30

FEATURE:
NAME/KEY: CDS
LOCATION: 34..1149
US-08-605-150A-7

Query Match 2.0%; Score 27; DB 3; Length 1390;
Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAATCTCGAG 1323
DB 1364 GTAAAAAATCTCGAG 1390

RESULT 6
US-09-370-838-17
Sequence 17, Application US/09370838
Patent No. 6444425
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamath, Roshon
APPLICANT: Secrest, Heather
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.475C1
CURRENT FILING DATE: 1999-08-09
CURRENT APPLICATION NUMBER: US/09/370,838
EARLIER FILING DATE: 1999-04-02
EARLIER APPLICATION NUMBER: US 09/285,323
NUMBER OF SEQ ID NOS: 289
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 317
TYPE: DNA
ORGANISM: Homo sapien
US-09-370-838-17

Query Match 2.0%; Score 26; DB 4; Length 317;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATCTCGAG 1323
DB 292 TAAAAAATCTCGAG 317

RESULT 7
US-08-946-026-10
Sequence 10, Application US/08946026
Patent No. 6034218
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Dillon, Davin C.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Mitcham, Jennifer L.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,026

FILING DATE: 07-OCT-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-946-026-10

Query Match 2.0%; Score 26; DB 3; Length 403;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATCTCGAG 1323
DB 378 TAAAAAATCTCGAG 403

RESULT 8
US-09-439-554-21
Sequence 21, Application US/09439554
Patent No. 6479733
GENERAL INFORMATION:
APPLICANT: Rafalecki, Jan Antoni
APPLICANT: Ogell, Joan T.
APPLICANT: Sakai, Hajime
APPLICANT: Thorpe, Catherine J.
APPLICANT: Kinney, Anthony J.
APPLICANT: Famodu, Omolayo O.
TITLE OF INVENTION: STEROL METABOLISM ENZYMES
FILE REFERENCE: BB114 US NA
CURRENT APPLICATION NUMBER: US/09/439,554
CURRENT FILING DATE: 1999-11-12
EARLIER APPLICATION NUMBER: 60/108,351
EARLIER FILING DATE: 1998-NO. 6479733ember-13
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 21
LENGTH: 815
TYPE: DNA
ORGANISM: Oryza sativa
US-09-439-554-21

Query Match 2.0%; Score 26; DB 4; Length 815;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATCTCGAG 1323
DB 628 TAAAAAATCTCGAG 653

RESULT 9
US-09-489-847-60
Sequence 60, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489,847
CURRENT FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657

EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 60
LENGTH: 1126
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (21)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (35)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (49)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (99)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1012)
OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-60

Query Match
Best Local Similarity 100.0%; Score 26; DB 4; Length 1126;
Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;
Matches 26; Conservative 0;

Qy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
Db 1101 TAAAAAAAAAAAAAAAAAATCGAG 1126

RESULT 10
US-09-674-529B-7
Sequence 7, Application US/09674529B
Patent No. 6551807
GENERAL INFORMATION:
APPLICANT: Cunningham, Francis X
TITLE OF INVENTION: CAROTENOID KETOYLASE GENES AND GENE PRODUCTS, PRODUCTION OF
FILE REFERENCE: 108172-00023
CURRENT APPLICATION NUMBER: US/09/674,529B
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: PCT/US99/10455
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,460
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 1141
TYPE: DNA
ORGANISM: Adonis aestivallis
US-09-674-529B-7

Query Match
Best Local Similarity 100.0%; Score 26; DB 4; Length 1141;
Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;
Matches 26; Conservative 0;

Qy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
Db 1116 TAAAAAAAAAAAAAAAAAATCGAG 1141

RESULT 11
US-08-747-562-11
Sequence 11, Application US/08747562
Patent No. 6579697
GENERAL INFORMATION:
APPLICANT: WALLACH, David
APPLICANT: BOLDIN, Mark
APPLICANT: METT, Igor
APPLICANT: VAPOLOMEYEV, Eugene
TITLE OF INVENTION: MODULATOR OF TNF/NGF SUPERFAMILY RECEPTORS
TITLE OF INVENTION: AND SOLUBLE OLIGOMERIC TNF/NGF SUPERFAMILY RECEPTORS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,562
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05854
FILING DATE: 11-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 109,632
FILING DATE: 11-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 111,125
FILING DATE: 02-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=15A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-747-562-11

Query Match
Best Local Similarity 100.0%; Score 26; DB 4; Length 1153;
Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;
Matches 26; Conservative 0;

Qy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
Db 1128 TAAAAAAAAAAAAAAAAAATCGAG 1153

RESULT 12
US-09-370-838-27
Sequence 27, Application US/09370838
Patent No. 6444425
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamach, Roadoh

APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
FILE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.475c1
CURRENT APPLICATION NUMBER: US/09/370,838
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/285,323
EARLIER FILING DATE: 1999-04-02
NUMBER OF SEQ ID NOS: 289
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 27
LENGTH: 1331
TYPE: DNA
ORGANISM: Homo sapien
US-09-370-838-27

Query Match 2.0%; Score 26; DB 4; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1306 TAAAAAAAAAAAAAAAACTCGAG 1331

RESULT 13

US-08-415-593-44
Sequence 44, Application US/08415593
Patent No. 5912140
Patent No. 5912140 5776726
GENERAL INFORMATION:
APPLICANT: Whoriskey, Susan K.
APPLICANT: Quinn, Cheryl L.
APPLICANT: Tao, Miatun
APPLICANT: Polietz-Vitz, Karen I.
APPLICANT: Schimmel, Paul R.
TITLE OF INVENTION: Recombinant Pneumococci Antinocyl rRNA
TITLE OF INVENTION: Synthetase Genes, Tester Strains and Assays
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA
APPLICATION NUMBER: US/08/415,593
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CP194-09
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-862-9540
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1352 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 101..1210
US-08-415-593-44

Query Match 2.0%; Score 26; DB 2; Length 1352;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1327 TAAAAAAAAAAAAAAAACTCGAG 1352

RESULT 14

US-09-149-476-207
Sequence 207, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 166 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596

EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 2.0%; Score 26; DB 4; Length 1354;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
|||||

Db 1287 TAAAAAAAAAAAACTGAG 1312

RESULT 15
US-09-149-476-23
Sequence 23, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149, 476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 2.0%; Score 26; DB 4; Length 1486;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATCTGAG 1323
DB 1423 TAAAAAATCTGAG 1448

RESULT 16
US-09-149-476-97
Sequence 97, Application US/09149476
Patent No. 6420526

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569

EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 2.0%; Score 26; DB 4; Length 1631;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
DB 1587 TAAAAAAAAAAAAAAAACTCGAG 1612

RESULT 17
US-09-205-258-175
Sequence 175, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 175
LENGTH: 2379
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (44)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1881)
OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-175
Query Match 2.0%; Score 26; DB 4; Length 2379;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 2320 TAAAAAAAAAAAAAAAACTCGAG 2345
RESULT 18
US-09-086-912-1
Sequence 1, Application US/09086912
Patent No. 6166192
GENERAL INFORMATION:
APPLICANT: Bruce M. Spiegelman, Pere Puigserver and Zhidan Wu
TITLE OF INVENTION: PGC-1, A No. 616619261 Brown Fat PPAR(SYMBOL
TITLE OF INVENTION: 103 {f "Symbol"} Coactivator
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHYE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/086,912
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/048,107
FILING DATE: 30-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy B.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-023
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3066 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 92..2482
US-09-086-912-1

Query Match 2.0%; Score 26; DB 3; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATACTCGAG 1323
DB 3009 TAAAAAATACTCGAG 3034

RESULT 19
US-09-203-453-1
Sequence 1, Application US/09203453
Patent No. 6426411
GENERAL INFORMATION:
APPLICANT: Spiegelman, Bruce M.; Puliserver, Pere; Wu, Zhidan and Adelmant, Guillaume
TITLE OF INVENTION: PGC-1, A NOVEL BROWN FAT PAR(SYMBOL 103 \F "Symbol") COACTIVATOR
FILE REFERENCE: DFN-023CP
CURRENT APPLICATION NUMBER: US/09/203,453
CURRENT FILING DATE: 1998-12-01
EARLIER APPLICATION NUMBER: 09/086,912
EARLIER FILING DATE: 1998-05-29
EARLIER APPLICATION NUMBER: 60/048,107
EARLIER FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3066
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (92)..(2482)
US-09-203-453-1

Query Match 2.0%; Score 26; DB 4; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATACTCGAG 1323
DB 3009 TAAAAAATACTCGAG 3034

RESULT 20
US-09-900-236-1
Sequence 1, Application US/09900236
Patent No. 6525178
GENERAL INFORMATION:
APPLICANT: Spiegelman, Bruce M.; Puliserver, Pere; Wu, Zhidan and
APPLICANT: Adelmant, Guillaume
TITLE OF INVENTION: PGC-1, A NOVEL BROWN FAT PAR(SYMBOL 103 \F "Symbol")
FILE REFERENCE: DFN-023CP
CURRENT APPLICATION NUMBER: US/09/900,236
CURRENT FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 09/203,453
PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: 60/048,107
PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3066
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (92)..(2482)
US-09-900-236-1

Query Match 2.0%; Score 26; DB 4; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATACTCGAG 1323
DB 3009 TAAAAAATACTCGAG 3034

RESULT 21
US-08-928-941D-28
Sequence 28, Application US/08928941D
Patent No. 6180763
GENERAL INFORMATION:
APPLICANT: Hirai, Hiroshi
TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,941D
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-002 N CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-343-1684
TELEFAX: 201-487-5800
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 3767 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-928-941D-28

Query Match 2.0%; Score 26; DB 3; Length 3767;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAATACTCGAG 1323
DB 3742 TAAAAAATACTCGAG 3767

RESULT 22
US-08-928-941D-30
Sequence 30, Application US/08928941D
Patent No. 6180763
GENERAL INFORMATION:
APPLICANT: Hirai, Hiroshi
APPLICANT: Sherr, Charles
TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES

TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,941D
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-002 N CIP
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 3767 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-928-941D-30
Query Match 2.0%; Score 26; DB 3; Length 3767;
Best Local Similarity 92.3%; Pred. No. 0.13;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAAAATCTCGAG 1323
DB 3742 UAAAAAAAAAAAAAAAAAATCTCGAG 3767
RESULT 23
US-09-280-590A-28
Sequence 28, Application US/09280590A
Patent No. 6303772
GENERAL INFORMATION:
APPLICANT: Hirai, Hiroshi
Sherri, Charles
Inoue, Kazushi
Bodner, Sarah M.
TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
THEREOF
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/280,590A
FILING DATE: 29-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 3767 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-280-590A-28
Query Match 2.0%; Score 26; DB 4; Length 3767;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1298 TAAAAAAAAAAAAAAAAAATCTCGAG 1323
DB 3742 TAAAAAAAAAAAAAAAAAATCTCGAG 3767
RESULT 24
US-09-280-590A-30
Sequence 30, Application US/09280590A
Patent No. 6303772
GENERAL INFORMATION:
APPLICANT: Hirai, Hiroshi
Sherri, Charles
Inoue, Kazushi
Bodner, Sarah M.
TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
THEREOF
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/280,590A
FILING DATE: 29-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 3767 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```
/
/
/ MOLECULE TYPE: linear
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-280-590A-30

Query Match
Best Local Similarity 92.3%; Score 26; DB 4; Length 3767;
Matches 24; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
Db 3742 UAAAAAAAAAAAAAAAAAATCGAG 3767

RESULT 25
US-09-023-905A-3
; Sequence 3, Application US/09023905A
; Patent No. 6475778
; GENERAL INFORMATION:
; APPLICANT: Roderes, Thomas M.
; APPLICANT: King, Frederick J.
; APPLICANT: Harris, David P.
; APPLICANT: Hu, Erding
; APPLICANT: Spiegelman, Bruce
; APPLICANT: Chan, Joanne
; TITLE OF INVENTION: Differentiation Enhancing Factors and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: DFN-021
; CURRENT APPLICATION NUMBER: US/09/023,905A
; CURRENT FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/038,191
; PRIOR FILING DATE: 1997-02-14
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4382
; TYPE: DNA
; ORGANISM: Danilo rerio
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (351)...(3803)
US-09-023-905A-3

Query Match
Best Local Similarity 100.0%; Score 26; DB 4; Length 4382;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
Db 4303 TAAAAAAAAAAAAAAAAAATCGAG 4328

RESULT 26
US-09-454-704A-2/C
; Sequence 2, Application US/09454704A
; Patent No. 6274321
; GENERAL INFORMATION:
; APPLICANT: Blumberg, Bruce
; APPLICANT: High Throughput Functional Screening of
; TITLE OF INVENTION: CDNA
; FILE REFERENCE: P-UC 3662
; CURRENT APPLICATION NUMBER: US/09/454,704A
; CURRENT FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
/
/ OTHER INFORMATION: Primer
US-09-454-704A-2

Query Match
Best Local Similarity 1.9%; Score 25; DB 3; Length 38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1299 AAAAAAAAAAAAAAAAAAATCGAG 1323
Db 37 AAAAAAAAAAAAAAAAAAATCGAG 13

RESULT 27
US-09-454-704A-12
; Sequence 12, Application US/09454704A
; Patent No. 6274321
; GENERAL INFORMATION:
; APPLICANT: Blumberg, Bruce
; TITLE OF INVENTION: High Throughput Functional Screening of
; TITLE OF INVENTION: CDNA
; FILE REFERENCE: P-UC 3662
; CURRENT APPLICATION NUMBER: US/09/454,704A
; CURRENT FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: double stranded CDNA
US-09-454-704A-12

Query Match
Best Local Similarity 1.9%; Score 25; DB 3; Length 38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1299 AAAAAAAAAAAAAAAAAAATCGAG 1323
Db 2 AAAAAAAAAAAAAAAAAAATCGAG 26

RESULT 28
US-09-454-704A-14
; Sequence 14, Application US/09454704A
; Patent No. 6274321
; GENERAL INFORMATION:
; APPLICANT: Blumberg, Bruce
; TITLE OF INVENTION: High Throughput Functional Screening of
; TITLE OF INVENTION: CDNA
; FILE REFERENCE: P-UC 3662
; CURRENT APPLICATION NUMBER: US/09/454,704A
; CURRENT FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CDNA
US-09-454-704A-14

Query Match
Best Local Similarity 1.9%; Score 25; DB 3; Length 42;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1299 AAAAAAAAAAAAAAAAAAATCGAG 1323
Db 2 AAAAAAAAAAAAAAAAAAATCGAG 26

RESULT 29
```

```

US-08-594-031-103
; Sequence 103, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-594-031-103

Query Match          1.9%; Score 25; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 138 AAAAAAAAAAAAAAAAACTCGAG 162

RESULT 30
; Sequence 3, Application US/08341568
; Patent No. 5661021
; GENERAL INFORMATION:
; APPLICANT: Buchert, Johanna
; APPLICANT: Silka-aho, Matti
; APPLICANT: Viikari, Liisa
; APPLICANT: Penttila, Merja
; APPLICANT: Saloheimo, Anu
; TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
; TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching 11g
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch

```

```

; STREET: PO Box 747
; CITY: Falls Church
; STATE: VA
; COUNTRY: US
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/341,568
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr, Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 365-262P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Trichoderma reesei
; STRAIN: QM9414
; US-08-341-568-3

Query Match          1.9%; Score 25; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 265 AAAAAAAAAAAAAAAAACTCGAG 289

RESULT 31
; Sequence 3, Application US/08911020
; Patent No. 5834047
; GENERAL INFORMATION:
; APPLICANT: Buchert, Johanna
; APPLICANT: Silka-aho, Matti
; APPLICANT: Viikari, Liisa
; APPLICANT: Penttila, Merja
; APPLICANT: Saloheimo, Anu
; TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
; TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: PO Box 747
; CITY: Falls Church
; STATE: VA
; COUNTRY: US
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,020
; FILING DATE: 13-AUG-1997

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/341,568
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr, Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 365-262P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei
STRAIN: QM9414
US-08-911-020-3

Query Match 1.9%; Score 25; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 265 AAAAAAAAAAAAAAAAACTCGAG 289

RESULT 32
US-09-091-725-28
Sequence 28, Application US/09091725
Patent No. 6329141
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Improved methods for transforming Phaffia
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster LLP
STREET: 2000 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,725
FILING DATE: 23-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95203620.0
FILING DATE: 22-DEC-1995
APPLICATION NUMBER: EP 96200943.7
FILING DATE: 11-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: E. Victor Donahue
REGISTRATION NUMBER: 35,492
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA

HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Phaffia rhodozyma
FEATURE:
NAME/KEY: CDS
LOCATION: 3..227
OTHER INFORMATION: /product= "PRCDNA18"
US-09-091-725-28

Query Match 1.9%; Score 25; DB 4; Length 307;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 265 AAAAAAAAAAAAAAAAACTCGAG 289

RESULT 33
US-08-341-568-5
Sequence 5, Application US/08341568
Patent No. 5651021
GENERAL INFORMATION:
APPLICANT: Buchert, Johanna
APPLICANT: Silka-aho, Matti
APPLICANT: Viikari, Liisa
APPLICANT: Penttila, Merja
APPLICANT: Saloheimo, Anu
APPLICANT: Marjatta, Rannu
TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching lign
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch and Birch
STREET: PO Box 747
CITY: Falls Church
STATE: VA
COUNTRY: US
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/341,568
FILING DATE: 22-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr, Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 365-262P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 319 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei
STRAIN: QM9414
US-08-341-568-5

Query Match 1.9%; Score 25; DB 1; Length 319;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
Db 295 AAAAAAAAAAAAAAAAACTCGAG 319

RESULT 34

US-08-911-020-5
; Sequence 5, Application US/08911020
; Patent No. 5854047
; GENERAL INFORMATION:
; APPLICANT: Buchert, Johanna
; APPLICANT: Silka-aho, Matti
; APPLICANT: Vilkar, Liisa
; APPLICANT: Penttila, Merja
; APPLICANT: Saloheimo, Anu
; APPLICANT: Marjatta, Ranna
; TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
; TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching
; TITLE OF INVENTION: lignocellulosic pulps
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: PO Box 747
; CITY: Falls Church
; STATE: VA
; COUNTRY: US
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911.020
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/341,568
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr, Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 365-262P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Trichoderma reesei
; STRAIN: QM9414
; US-08-911-020-5

Query Match 1.9%; Score 25; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
Db 295 AAAAAAAAAAAAAAAAACTCGAG 319

RESULT 35
US-08-946-026-35
; Sequence 35, Application US/08946026
; Patent No. 6034218

GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.
; APPLICANT: Dillon, Devin C.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Mitcham, Jennifer L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,026
; FILING DATE: 07-OCT-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.424C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-946-026-35

Query Match 1.9%; Score 25; DB 3; Length 333;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
Db 309 AAAAAAAAAAAAAAAAACTCGAG 333

RESULT 36
US-08-888-077A-39/C
; Sequence 39, Application US/08888077A
; Patent No. 6020143
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLTZ & MENTLIK
; STREET: 600 SOUTH AVENUE WEST
; CITY: WESTFIELD
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,077A

```
/
/ FILING DATE: 03-JUL-1997
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/592,541
/ FILING DATE: 26-JAN-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: PALISI, THOMAS M
/ REGISTRATION NUMBER: 36,629
/ REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP IV
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (908) 654-5000
/ TELEFAX: (908) 654-7866
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 340 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 1..350
/ OTHER INFORMATION: /note="PS1EX10/24-1"
/
US-08-888-077A-39
/
Query Match 1.9%; Score 25; DB 3; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 25 AAAAAAAAAAAAAAAAACTCGAG 1

RESULT 37
US-08-888-077A-37/C
/ Sequence 37, Application US/08888077A
/ Patent No. 6020143
/ GENERAL INFORMATION:
/ APPLICANT: ST. GEORGE-HYSLOP, PETER H
/ APPLICANT: ROMMENS, JOHANNA M
/ APPLICANT: FRASER, PAUL E
/ TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
/ TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.
/ NUMBER OF SEQUENCES: 41
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLTZ & MENTLIK
/ STREET: 600 SOUTH AVENUE WEST
/ CITY: WESTFIELD
/ STATE: NJ
/ COUNTRY: USA
/ ZIP: 07090-1497
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/888,077A
/ FILING DATE: 03-JUL-1997
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/592,541
/ FILING DATE: 26-JAN-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: PALISI, THOMAS M
/ REGISTRATION NUMBER: 36,629
/ REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP IV
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (908) 654-7866
/ TELEFAX: (908) 654-5000
/ INFORMATION FOR SEQ ID NO: 37:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 360 base pairs
```

```
/
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 1..360
/ OTHER INFORMATION: /note="PS1LY2HEX10-17-1"
/
US-08-888-077A-37
/
Query Match 1.9%; Score 25; DB 3; Length 360;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 25 AAAAAAAAAAAAAAAAACTCGAG 1

RESULT 38
US-08-946-026-23
/ Sequence 23, Application US/08946026
/ Patent No. 6034218
/ GENERAL INFORMATION:
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Dillon, Devin C.
/ APPLICANT: Twardzik, Daniel R.
/ APPLICANT: Mitcham, Jennifer L.
/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
/ TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
/ NUMBER OF SEQUENCES: 59
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: SEED and BERRY LLP
/ STREET: 6300 Columbia Center, 701 Fifth Avenue
/ CITY: Seattle
/ STATE: Washington
/ COUNTRY: USA
/ ZIP: 98104-7092
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/946,026
/ FILING DATE: 07-OCT-1997
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MAKI, David J.
/ REGISTRATION NUMBER: 31,392
/ REFERENCE/DOCKET NUMBER: 210121.424C1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 622-4900
/ TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 375 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
US-08-946-026-23
/
Query Match 1.9%; Score 25; DB 3; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 351 AAAAAAAAAAAAAAAAACTCGAG 375

RESULT 39
US-09-220-132-128/C
/ Sequence 128, Application US/09220132
```

Patent No. 6506607
GENERAL INFORMATION:
APPLICANT: Shvjan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
FILE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 0734-074001
CURRENT APPLICATION NUMBER: US/09/220,132
CURRENT FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 128
LENGTH: 376
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(376)
OTHER INFORMATION: n = A,T,C or G
US-09-220-132-128

Query Match 1.9%; Score 25; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 61 AAAAAAAAAAAAAAAAACTCGAG 37

RESULT 40
US-08-946-026-41
Sequence 41, Application US/08946026
Patent No. 6034218
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Dillon, Davin C.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Mitcham, Jennifer L.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,026
FILING DATE: 07-OCT-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121,424C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
US-08-946-026-41

Query Match 1.9%; Score 25; DB 3; Length 377;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 353 AAAAAAAAAAAAAAAAACTCGAG 377

RESULT 41
US-09-370-838-164
Sequence 164, Application US/09370838
Patent No. 6444425
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamach, Roadoh
APPLICANT: Secretat, Heather
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.475C1
CURRENT APPLICATION NUMBER: US/09/370,838
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/285,323
EARLIER FILING DATE: 1999-04-02
NUMBER OF SEQ ID NOS: 289
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 164
LENGTH: 395
TYPE: DNA
ORGANISM: Homo sapien
US-09-370-838-164

Query Match 1.9%; Score 25; DB 4; Length 395;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
DB 371 AAAAAAAAAAAAAAAAACTCGAG 395

RESULT 42
US-08-630-822A-92
Sequence 92, Application US/08630822A
Patent No. 5840695
GENERAL INFORMATION:
APPLICANT: FRANK, GLENN R.
APPLICANT: HUNTER, SHIRLEY WU
APPLICANT: WALLENFELS, LYNDIA
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,822A
FILING DATE: 11-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: CONNELL, GARY J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 490 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-630-822A-92

Query Match 1.9%; Score 25; DB 2; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
Db 462 AAAAAAAAAAAAAAAAACTCGAG 486

RESULT 43
US-09-005-069-92
Sequence 92, Application US/09005069
Patent No. 5933470
GENERAL INFORMATION:
APPLICANT: FRANK, GLENN R.
APPLICANT: HUNTER, SHIRLEY WU
APPLICANT: WALLENFELS, LYNDA
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,069
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,822
FILING DATE: 11-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: CONNELL, GARY J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 490 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-005-069-92

Query Match 1.9%; Score 25; DB 2; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
Db 462 AAAAAAAAAAAAAAAAACTCGAG 486

RESULT 44
US-09-171-156A-41
Sequence 41, Application US/09171156A
Patent No. 6368846
GENERAL INFORMATION:
APPLICANT: Hunter, Shirley Wu
APPLICANT: Sim, Gek-kee
APPLICANT: Weber, Eric R.
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: SHERIDAN ROSS P.C.
STREET: 1560 BROADWAY, SUITE 1200
CITY: DENVER
STATE: CO
COUNTRY: U.S.A.
ZIP: 80202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,156A
FILING DATE: 04-Mar-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 490 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-171-156A-41

Query Match 1.9%; Score 25; DB 4; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
Db 462 AAAAAAAAAAAAAAAAACTCGAG 486

RESULT 45
US-09-004-730A-41
Sequence 41, Application US/09004730A
Patent No. 6485968
GENERAL INFORMATION:
APPLICANT: Weber, Eric
APPLICANT: Wu Hunter, Shirley
APPLICANT: Sim, Gek-kee
APPLICANT: Frank, Glenn
APPLICANT: Wallenfels, Lynda
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND APPARATUS TO COLLECT SUCH
FILE REFERENCE: 2618-17-C5-PUS-1
CURRENT APPLICATION NUMBER: US/09/004,730A
CURRENT FILING DATE: 1998-01-08

PRIOR APPLICATION NUMBER: PCT/97US/18669
PRIOR FILING DATE: 1997-10-15
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn version 3.0
SEQ ID NO 41
LENGTH: 490
TYPE: DNA
ORGANISM: Ctenocephalides felis
US-09-004-730A-41

Query Match 1.9%; Score 25; DB 4; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 462 AAAAAAAAAAAAAAAAAAACTCGAG 486

RESULT 46
US-08-981-799A-41
Sequence 41, Application US/08981799A
Patent No. 6576238
GENERAL INFORMATION:
APPLICANT: Weber, Eric
APPLICANT: Wu Hunter, Shirley
APPLICANT: Sim, Gek-kee
APPLICANT: Frank, Glenn
APPLICANT: Wallenfels, Lynda
TITLE OF INVENTION: "NOVEL ECTOPARASITE SALIVA PROTEINS AND APPARATUS TO COLLECT SU
FILE REFERENCE: 2618-17-CS-PUS
CURRENT APPLICATION NUMBER: US/08/981,799A
CURRENT FILING DATE: 1998-08-27
PRIOR APPLICATION NUMBER: PCT/97/18669
PRIOR FILING DATE: 1997-10-15
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn version 3.0
SEQ ID NO 41
LENGTH: 490
TYPE: DNA
ORGANISM: Ctenocephalides felis
US-08-981-799A-41

Query Match 1.9%; Score 25; DB 4; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 462 AAAAAAAAAAAAAAAAAAACTCGAG 486

RESULT 47
US-09-370-838-263
Sequence 263, Application US/09370838
Patent No. 6444425
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamach, Roadah
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
FILE REFERENCE: 210121.475C1
CURRENT APPLICATION NUMBER: US/09/370,838
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/285,323
EARLIER FILING DATE: 1999-04-02
NUMBER OF SEQ ID NOS: 289
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 263
LENGTH: 506

TYPE: DNA
ORGANISM: Homo sapiens
US-09-370-838-263

Query Match 1.9%; Score 25; DB 4; Length 506;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 482 AAAAAAAAAAAAAAAAAAACTCGAG 506

RESULT 48
US-08-341-568-1
Sequence 1, Application US/08341568
Patent No. 5661021
GENERAL INFORMATION:
APPLICANT: Buchert, Johanna
APPLICANT: Silka-aho, Matti
APPLICANT: Vilkar, Liisa
APPLICANT: Penttila, Merja
APPLICANT: Saloheimo, Anu
APPLICANT: Marjatta, Rannu
TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching lign
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: Birch, Stewart, Kolaach and Birch
STREET: PO Box 747
CITY: Falls Church
STATE: VA
COUNTRY: US
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/341,568
FILING DATE: 22-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr, Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 365-262P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 536, base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: trichoderma reesei
STRAIN: QM9414
US-08-341-568-1

Query Match 1.9%; Score 25; DB 1; Length 536;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 AAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 512 AAAAAAAAAAAAAAAAAAACTCGAG 536

RESULT 49
US-08-911-020-1
Sequence 1, Application US/08911020
Patent No. 5854047
GENERAL INFORMATION:
APPLICANT: Buchert, Johanna
APPLICANT: Silka-aho, Matti
APPLICANT: Vilkari, Liisa
APPLICANT: Penttila, Merja
APPLICANT: Seicheimo, Anu
APPLICANT: Marjatta, Rana
TITLE OF INVENTION: Mannanase enzymes, genes coding for them
TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolaesch and Birch
STREET: PO Box 747
CITY: Falls Church
STATE: VA
COUNTRY: US
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,020
FILING DATE: 13-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/341,568
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr, Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 365-262P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 536 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei
STRAIN: QM9414
US-08-911-020-1

Query Match 1.9%; Score 25; DB 2; Length 536;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
Db 512 AAAAAAAAAAAAAAAAACTCGAG 536

RESULT 50
US-09-370-838-242
Sequence 242, Application US/09370838
Patent No. 6444425
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Monamach, Roshoh
APPLICANT: Secretet, Heather
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF

TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.475C1
CURRENT APPLICATION NUMBER: US/09/370,838
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/285,323
EARLIER FILING DATE: 1999-04-02
NUMBER OF SEQ ID NOS: 289
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 242
LENGTH: 556
TYPE: DNA
ORGANISM: Homo sapiens
US-09-370-838-242

Query Match 1.9%; Score 25; DB 4; Length 556;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1299 AAAAAAAAAAAAAAAAACTCGAG 1323
Db 532 AAAAAAAAAAAAAAAAACTCGAG 556

Search completed: January 28, 2004, 17:27:05
Job time : 115 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 15:56:05 ; Search time 440 Seconds
(without alignments)
10753.797 Million cell updates/sec

Title: US-09-881-556a-3
Perfect score: 1323
Sequence: 1 gattccgcgcagcagattttt.....aaaaaaaaaactcgag 1323

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2356869 seqs, 1788235258 residues

Word size : 15

Total number of hits satisfying chosen parameters: 243227

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Published Applications NA.*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2.2	538	11	US-09-918-995-22177	Sequence 22177, A
2	2.2	1049	11	US-09-809-391-58	Sequence 58, App1
3	2.2	1049	13	US-09-882-171-58	Sequence 58, App1
4	2.2	1108	12	US-10-264-237-619	Sequence 619, App
5	2.2	1198	11	US-09-776-724A-60	Sequence 60, App
6	2.1	743	15	US-10-102-806-109	Sequence 109, App
7	2.1	3150	10	US-09-925-300-83	Sequence 83, App1
8	2.1	7394	13	US-09-814-353-20112	Sequence 20112, A
9	2.0	480	11	US-09-918-995-24814	Sequence 24814, A
10	2.0	708	9	US-09-764-869-465	Sequence 465, App
11	2.0	708	12	US-10-327-577-465	Sequence 465, App
12	2.0	708	15	US-10-091-504-465	Sequence 465, App
13	2.0	952	13	US-09-814-353-19712	Sequence 19712, A
14	2.0	1061	10	US-09-738-973-149	Sequence 149, App
15	2.0	1061	10	US-09-854-133-149	Sequence 149, App

16	2.0	1061	15	US-10-144-649A-149	Sequence 149, App
17	2.0	1227	12	US-10-264-237-840	Sequence 840, App
18	2.0	1553	27	US-09-925-301-305	Sequence 305, App
19	2.0	309	11	US-09-776-724A-54	Sequence 54, App1
20	2.0	317	10	US-09-738-973-17	Sequence 17, App1
21	2.0	317	10	US-09-854-133-17	Sequence 17, App1
22	2.0	317	15	US-10-144-649A-17	Sequence 17, App1
23	2.0	406	10	US-09-783-590-7936	Sequence 7936, App
24	2.0	478	11	US-09-918-995-19475	Sequence 19475, A
25	2.0	475	11	US-09-918-995-22561	Sequence 22561, A
26	2.0	502	11	US-09-918-995-23066	Sequence 23066, A
27	2.0	513	9	US-09-918-995-19781	Sequence 19781, A
28	2.0	519	11	US-09-764-887-134	Sequence 134, App
29	2.0	519	15	US-10-073-961-134	Sequence 134, App
30	2.0	591	15	US-10-106-698-1187	Sequence 1187, App
31	2.0	646	15	US-10-106-698-517	Sequence 517, App
32	2.0	682	11	US-09-823-129-16	Sequence 16, App1
33	2.0	695	13	US-10-342-224-3	Sequence 3, App1
34	2.0	786	10	US-09-966-881-26	Sequence 26, App1
35	2.0	825	13	US-10-027-632-155629	Sequence 155629, App1
36	2.0	825	14	US-10-027-632-155629	Sequence 155629, App1
37	2.0	1076	12	US-10-264-237-135	Sequence 135, App
38	2.0	1108	10	US-09-872-523-4	Sequence 4, App1
39	2.0	1108	10	US-09-872-523-48	Sequence 48, App1
40	2.0	1143	13	US-09-814-353-21392	Sequence 21392, A
41	2.0	1131	10	US-09-738-973-27	Sequence 27, App1
42	2.0	1131	10	US-09-854-133-27	Sequence 27, App1
43	2.0	1131	15	US-10-144-649A-27	Sequence 27, App1
44	2.0	1134	15	US-10-231-417-129	Sequence 129, App
45	2.0	1150	9	US-09-925-299-82	Sequence 82, App1
46	2.0	1150	11	US-09-925-299-82	Sequence 82, App1
47	2.0	1150	15	US-10-106-698-2096	Sequence 2096, App
48	2.0	1154	11	US-09-809-391-207	Sequence 207, App
49	2.0	1154	13	US-09-882-171-207	Sequence 207, App
50	2.0	1133	13	US-10-050-704-52	Sequence 52, App1
51	2.0	1486	11	US-09-809-391-23	Sequence 23, App1
52	2.0	1486	11	US-09-882-171-23	Sequence 23, App1
53	2.0	1576	11	US-09-892-877-48	Sequence 48, App1
54	2.0	1576	11	US-09-948-783-108	Sequence 108, App
55	2.0	1531	11	US-09-809-391-97	Sequence 97, App1
56	2.0	1531	11	US-09-882-171-97	Sequence 97, App1
57	2.0	1536	11	US-09-892-877-16	Sequence 16, App1
58	2.0	1536	11	US-09-948-783-16	Sequence 16, App1
59	2.0	1811	15	US-10-198-846-11524	Sequence 11524, A
60	2.0	1842	10	US-09-834-975-783	Sequence 783, App
61	2.0	1906	13	US-10-412-000-1	Sequence 1, App1
62	2.0	1906	14	US-10-021-657-1	Sequence 1, App1
63	2.0	2007	15	US-10-106-698-1957	Sequence 1957, App
64	2.0	2185	12	US-10-264-049-633	Sequence 633, App
65	2.0	2254	9	US-09-800-729-52	Sequence 52, App1
66	2.0	2379	13	US-09-933-767-175	Sequence 175, App
67	2.0	2379	15	US-10-023-282-175	Sequence 175, App
68	2.0	2674	7	US-10-097-065-15	Sequence 15, App1
69	2.0	3066	9	US-08-731-499-12	Sequence 12, App1
70	2.0	3066	9	US-09-900-236-1	Sequence 1, App1
71	2.0	3066	12	US-10-439-799-1	Sequence 1, App1
72	2.0	3066	15	US-10-066-443-1	Sequence 1, App1
73	2.0	3066	15	US-10-256-889-1	Sequence 1, App1
74	2.0	3315	15	US-10-198-846-9945	Sequence 9945, App
75	2.0	3367	11	US-09-892-338-28	Sequence 28, App1
76	2.0	3367	11	US-09-892-338-10	Sequence 30, App1
77	2.0	7474	11	US-09-945-917-26	Sequence 26, App1
78	2.0	7474	13	US-09-845-917A-26	Sequence 26, App1
79	2.0	10443	11	US-09-945-917-25	Sequence 25, App1
80	2.0	10443	13	US-09-845-917A-25	Sequence 25, App1
81	1.9	28	13	US-10-354-152-7	Sequence 7, App1
82	1.9	38	13	US-10-371-077-13	Sequence 13, App1
83	1.9	42	13	US-10-277-969-10	Sequence 10, App1
84	1.9	42	13	US-10-277-969-18	Sequence 18, App1
85	1.9	42	15	US-10-200-148-10	Sequence 10, App1
86	1.9	42	15	US-10-200-148-18	Sequence 18, App1
87	1.9	44	14	US-10-094-163-9	Sequence 9, App1
88	1.9	54	10	US-09-783-550-3459	Sequence 3459, App

```
C 89 25 1.9 59 10 US-09-783-590-2044 Sequence 2044, Ap
C 90 25 1.9 70 10 US-09-783-590-4790 Sequence 4790, Ap
C 91 25 1.9 82 10 US-09-783-590-1257 Sequence 1257, Ap
C 92 25 1.9 87 10 US-09-783-590-2629 Sequence 2629, Ap
C 93 25 1.9 180 9 US-09-925-299-501 Sequence 501, App
C 94 25 1.9 180 11 US-09-925-299-501 Sequence 501, App
C 95 25 1.9 286 9 US-09-833-790-16 Sequence 16, Appl
C 96 25 1.9 298 9 US-09-833-790-14 Sequence 14, Appl
C 97 25 1.9 321 9 US-09-833-790-26 Sequence 26, Appl
C 98 25 1.9 349 10 US-09-783-590-12129 Sequence 12129, A
C 99 25 1.9 364 9 US-09-833-790-301 Sequence 301, App
C 100 25 1.9 364 10 US-09-854-133-631 Sequence 631, App
```

ALIGNMENTS

```
RESULT 1
US-09-918-995-22177
; Sequence 22177, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 22177
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(538)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-22177
```

```
Query Match 2.2%; Score 29; DB 11; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.0057;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 505 AAGTAAAAAAAAAAAAAAAAAACTCGAG 533
```

```
RESULT 2
US-09-809-391-58
; Sequence 58, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 1049
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-809-391-58
```

```
Query Match 2.2%; Score 29; DB 11; Length 1049;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 1012 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1040
```

```
RESULT 3
US-09-882-171-58
; Sequence 58, Application US/09882171
; Publication No. US20030175858A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/882,171
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/809,391
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,581
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,584
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,500
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,613
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,582
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,596
```

PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,612
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,632
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,601
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/043,580
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,568
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,314
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,569
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,311
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,671
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,674
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,669
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,312
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,313
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,672
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,315
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/048,974
 PRIOR FILING DATE: 1997-06-06
 PRIOR APPLICATION NUMBER: 60/056,886
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,877
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,889
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,893
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,630
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,878
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,662
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,872
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,882
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,637
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,903
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,888
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,879
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,880
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,894
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,911
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,636
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,874
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,910
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,864
 PRIOR FILING DATE: 1997-08-22

PRIOR APPLICATION NUMBER: 60/056,631
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,845
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,892
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/057,761
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/047,595
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,599
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,588
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,585
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,586
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,590
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,594
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,589
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,593
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/047,614
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/043,578
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/043,576
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/047,501
 PRIOR FILING DATE: 1997-05-23
 PRIOR APPLICATION NUMBER: 60/043,670
 PRIOR FILING DATE: 1997-04-11
 PRIOR APPLICATION NUMBER: 60/056,632
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,664
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,876
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,881
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,909
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,875
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,862
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,887
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/056,908
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/048,964
 PRIOR FILING DATE: 1997-06-06
 PRIOR APPLICATION NUMBER: 60/057,650
 PRIOR FILING DATE: 1997-09-05
 PRIOR APPLICATION NUMBER: 60/056,884
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: 60/057,669
 PRIOR FILING DATE: 1997-09-05

Query Match 2.2% Score 29; DB 13; Length 1049;
 Best Local Similarity 100.0%; Pred No. 0.006;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1295 AAGTAAAAAAAAAAAAAAAAAAGCTGAG 1323
 Db 1012 AAGTAAAAAAAAAAAAAAAAAAGCTGAG 1040

RESULT 4

```
US-10-264-237-619
; Sequence 619, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birtse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 619
; LENGTH: 1108
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1101)..(1101)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-619

Query Match          2.2%; Score 29; DB 12; Length 1108;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
Db      1041 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1069

RESULT 5
US-09-776-724A-60
; Sequence 60, Application US/09776724A
; Publication No. US20030050455A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 64 Human Secreted Proteins
; FILE REFERENCE: P2011
; CURRENT APPLICATION NUMBER: US/09/776,724A
; CURRENT FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/180,909
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 09/669,688
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: 09/229,982
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: PCT/US98/14613
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/052,661
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,872
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 109
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-806-109/c

Query Match          2.1%; Score 28; DB 15; Length 743;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,725
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,985
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,952
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,989
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,361
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,726
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,724
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,946
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,683
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 1198
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1189)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1191)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-776-724A-60
```

```
Query Match          2.2%; Score 29; DB 11; Length 1198;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
Db      1135 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1163
```

```
RESULT 6
US-10-102-806-109/c
; Sequence 109, Application US/10102806
; Publication No. US2003005421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 109
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-806-109
```

```
Query Match          2.1%; Score 28; DB 15; Length 743;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 28 AAGTAAAAAAAAAAAAAACTCGA 1

RESULT 7

US-09-925-300-83

Sequence 83, Application US/09925300

Patent No. US20020151681A1

GENERAL INFORMATION:

APPLICANT: Craig Roosen,

APPLICANT: Steve Ruben,

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA101

CURRENT APPLICATION NUMBER: US/09/925,300

CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05988

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 1890

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 83

LENGTH: 2150

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-300-83

Query Match

Best Local Similarity 2.1%; Score 28; DB 10; Length 2150;

Best Local Similarity 100.0%; Pred. No. 0.018; Mismatches 0; Indels 0; Gaps 0;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1295 AAGTAAAAAAAAAAAAAACTCGA 1322

Db 2123 AAGTAAAAAAAAAAAAAACTCGA 2150

RESULT 8

US-09-814-353-20112

Sequence 20112, Application US/09814353

Publication No. US20030165831A1

GENERAL INFORMATION:

APPLICANT: Lee, John

APPLICANT: Thompson, Pamela

APPLICANT: Lillie, James

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER

FILE REFERENCE: MRI-006B

CURRENT APPLICATION NUMBER: US/09/814,353

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: US 60/191,031

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: US 60/207,124

PRIOR FILING DATE: 2000-05-25

PRIOR APPLICATION NUMBER: US 60/211,940

PRIOR FILING DATE: 2000-06-15

PRIOR APPLICATION NUMBER: US 60/216,820

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: US 60/220,661

PRIOR FILING DATE: 2000-07-25

PRIOR APPLICATION NUMBER: US 60/257,672

PRIOR FILING DATE: 2000-12-21

NUMBER OF SEQ ID NOS: 22037

SOFTWARE: PatentIn Ver. 4.0

SEQ ID NO 20112

LENGTH: 3794

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..2, 3784, 3785, 3786, 3787, 3788, 3789, 3790, 3791, 3792,

LOCATION: 3793, 3794

OTHER INFORMATION: n = A,T,C or G

US-09-814-353-20112

Query Match

Best Local Similarity 2.1%; Score 28; DB 13; Length 3794;

Best Local Similarity 100.0%; Pred. No. 0.019; Mismatches 0; Indels 0; Gaps 0;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1295 AAGTAAAAAAAAAAAAAACTCGA 1322

Db 3756 AAGTAAAAAAAAAAAAAACTCGA 3783

RESULT 9

US-09-918-995-24814

Sequence 24814, Application US/09918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED FROM VARIOUS CDNA LIBRARIES

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

CURRENT FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235,076

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: PatentIn Ver. 3.0

SEQ ID NO 24814

LENGTH: 480

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)...(480)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-24814

Query Match

Best Local Similarity 2.0%; Score 27; DB 11; Length 480;

Best Local Similarity 100.0%; Pred. No. 0.048; Mismatches 0; Indels 0; Gaps 0;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1295 AAGTAAAAAAAAAAAAAACTCG 1321

Db 430 AAGTAAAAAAAAAAAAAACTCG 456

RESULT 10

US-09-764-869-465/C

Sequence 465, Application US/09764869

Patent No. US20020061521A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC007

CURRENT APPLICATION NUMBER: US/09/764,869

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 2442

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 465

LENGTH: 708

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (1)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE

LOCATION: (5)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE

LOCATION: (671)

OTHER INFORMATION: n equals a,t,g, or c

US-09-764-869-465


```
Query Match          2.0%; Score 27; DB 9; Length 708;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1297 GTAAAAAAAAAAAAAAAACTCGAG 1323
DB      51  GTAAAAAAAAAAAAAAAACTCGAG 25

RESULT 11
US-10-227-577-465/c
; Sequence 465, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 465
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (671)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-227-577-465

Query Match          2.0%; Score 27; DB 12; Length 708;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1297 GTAAAAAAAAAAAAAAAACTCGAG 1323
DB      51  GTAAAAAAAAAAAAAAAACTCGAG 25

RESULT 12
US-10-091-504-465/c
; Sequence 465, Application US/10091504
; Publication No. US2003005908A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 465
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (671)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-091-504-465

Query Match          2.0%; Score 27; DB 15; Length 708;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1297 GTAAAAAAAAAAAAAAAACTCGAG 1323
DB      51  GTAAAAAAAAAAAAAAAACTCGAG 25

RESULT 13
US-09-814-353-19712/c
; Sequence 19712, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19712
; LENGTH: 952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1_2_3_812_818_891_897_899_902_945_946_947_948,
; LOCATION: 949_950_951_952
; OTHER INFORMATION: n = A,T,C or G
; US-09-814-353-19712
```

Query Match 2.0%; Score 27; DB 13; Length 952;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 55 GTAAAAAAAAAAAAAAAAAACTCGAG 29

RESULT 14
US-09-738-973-149

; Sequence 149, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indictae, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738, 973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 149
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-738-973-149

Query Match 2.0%; Score 27; DB 10; Length 1061;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 1035 GTAAAAAAAAAAAAAAAAAACTCGAG 1061

RESULT 15

US-09-854-133-149
; Sequence 149, Application US/09854133
; Publication No. US2002018349A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854, 133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 149
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-149

Query Match 2.0%; Score 27; DB 10; Length 1061;
Best Local Similarity 100.0%; Pred. No. 0.051;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 1035 GTAAAAAAAAAAAAAAAAAACTCGAG 1061

RESULT 16
US-10-144-649A-149

; Sequence 149, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongrong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144, 649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 149
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-144-649A-149

Query Match 2.0%; Score 27; DB 15; Length 1061;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 1035 GTAAAAAAAAAAAAAAAAAACTCGAG 1061

RESULT 17

US-10-264-237-840
; Sequence 840, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4131P1
; CURRENT APPLICATION NUMBER: US/10/264, 237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205, 515
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 840
; LENGTH: 1427

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (378)..(378)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1256)..(1256)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1282)..(1282)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature

```
LOCATION: (1298)..(1298)
OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-840

Query Match
Best Local Similarity 100.0%; Score 27; DB 12; Length 1427;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1323
DB 1199 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1225

RESULT 18
US-09-925-301-305
Sequence 305, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 305
LENGTH: 1553
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-301-305

Query Match
Best Local Similarity 100.0%; Score 27; DB 9; Length 1553;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1297 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1323
DB 1527 GTAAAAAAAAAAAAAAAAAAGCTCGAG 1553

RESULT 19
US-09-776-724A-54
Sequence 54, Application US/0976724A
Publication No. US20030050455A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 64 Human Secreted Proteins
FILE REFERENCE: P2011
CURRENT APPLICATION NUMBER: US/09/776,724A
CURRENT FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 60/180,509
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 09/669,688
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: 09/229,982
PRIOR FILING DATE: 1999-01-14
PRIOR APPLICATION NUMBER: PCT/US98/14613
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/052,661
PRIOR FILING DATE: 1997-07-16
PRIOR APPLICATION NUMBER: 60/052,872
PRIOR FILING DATE: 1997-07-16
PRIOR APPLICATION NUMBER: 60/052,871
PRIOR FILING DATE: 1997-07-16
PRIOR APPLICATION NUMBER: 60/052,874
PRIOR FILING DATE: 1997-07-16
PRIOR APPLICATION NUMBER: 60/052,873
PRIOR FILING DATE: 1997-07-16
PRIOR APPLICATION NUMBER: 60/052,870
```

```
PRIOR FILING DATE: 1997-07-16
PRIOR APPLICATION NUMBER: 60/052,875
PRIOR FILING DATE: 1997-07-16
PRIOR APPLICATION NUMBER: 60/053,440
PRIOR FILING DATE: 1997-07-22
PRIOR APPLICATION NUMBER: 60/053,441
PRIOR FILING DATE: 1997-07-22
PRIOR APPLICATION NUMBER: 60/053,442
PRIOR FILING DATE: 1997-07-22
PRIOR APPLICATION NUMBER: 60/056,359
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,725
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,985
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,952
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,989
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/056,361
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,726
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,724
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,946
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,683
PRIOR FILING DATE: 1997-08-18
NUMBER OF SEQ ID NOS: 294
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 54
LENGTH: 309
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (301)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (305)
OTHER INFORMATION: n equals a,t,g, or c
US-09-776-724A-54

Query Match
Best Local Similarity 100.0%; Score 26; DB 11; Length 309;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
DB 240 TAAAAAAAAAAAAAAAAAAGCTCGAG 265

RESULT 20
US-09-738-973-17
Sequence 17, Application US/09738973
Patent No. US20020110563A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Henderson, Robert A.
APPLICANT: Lodes, Michael J.
APPLICANT: Fling, Steven P.
APPLICANT: Mohamath, Raodoh
APPLICANT: Algate, Paul A.
APPLICANT: Secrist, Heather
APPLICANT: Indrias, Carol Joseph
APPLICANT: Benson, Darin R.
APPLICANT: Elliot, Mark
APPLICANT: Mannon, Jane
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
FILE REFERENCE: 210121.475C9
TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
```

;; CURRENT APPLICATION NUMBER: US/09/738,973
;; CURRENT FILING DATE: 2000-12-14
;; NUMBER OF SEQ ID NOS: 587
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 17
;; LENGTH: 317
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-09-738-973-17

Query Match 2.0%; Score 26; DB 10; Length 317;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 292 TAAAAAAAAAAAAAAAAAACTCGAG 317

RESULT 21
US-09-854-133-17
;; Sequence 17, Application US/09854133
;; Publication No. US20020183499A1
;; GENERAL INFORMATION:
;; APPLICANT: Lodes, Michael J.
;; APPLICANT: Mohamath, Raedoh
;; APPLICANT: Henderson, Robert A.
;; APPLICANT: Benson, Darin R.
;; APPLICANT: Secrist, Heather
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.475C10
;; CURRENT APPLICATION NUMBER: US/09/854,133
;; CURRENT FILING DATE: 2001-05-11
;; NUMBER OF SEQ ID NOS: 735
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 17
;; LENGTH: 317
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-09-854-133-17

Query Match 2.0%; Score 26; DB 10; Length 317;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 292 TAAAAAAAAAAAAAAAAAACTCGAG 317

RESULT 22
US-10-144-649A-17
;; Sequence 17, Application US/10144649A
;; Publication No. US20030118599A1
;; GENERAL INFORMATION:
;; APPLICANT: Lodes, Michael J.
;; APPLICANT: Wang, Tonglong
;; APPLICANT: Fan, Liqun
;; APPLICANT: Algate, Paul A.
;; APPLICANT: McNeill, Patricia D.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.475C11
;; CURRENT APPLICATION NUMBER: US/10/144,649A
;; CURRENT FILING DATE: 2002-08-21
;; NUMBER OF SEQ ID NOS: 749
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 17
;; LENGTH: 317
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-144-649A-17

Query Match 2.0%; Score 26; DB 15; Length 317;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 292 TAAAAAAAAAAAAAAAAAACTCGAG 317

RESULT 23
US-09-783-590-7936/c
;; Sequence 7936, Application US/09783590
;; Patent No. US20020110850A1
;; GENERAL INFORMATION:
;; APPLICANT: Dillion, Patrick J.
;; APPLICANT: Haseltine, William A.
;; APPLICANT: Li, Haodong
;; APPLICANT: Rosen, Craig A.
;; APPLICANT: Ruben, Steven M.
;; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;; FILE REFERENCE: PO-16 2C1
;; CURRENT APPLICATION NUMBER: US/09/783,590
;; CURRENT FILING DATE: 2000-02-15
;; PRIOR APPLICATION NUMBER: 08/420,856
;; PRIOR FILING DATE: 1995-04-12
;; PRIOR APPLICATION NUMBER: 08/346,731
;; PRIOR FILING DATE: 1994-11-21
;; NUMBER OF SEQ ID NOS: 12485
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 7936
;; LENGTH: 406
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (9)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (130)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (158)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (187)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (218)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (248)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (256)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (262)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (283)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (289)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (328)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (329)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc feature
;; LOCATION: (338)

```
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (362)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (379)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (402)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (406)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-7936
```

```
Query Match          2.0%; Score 26; DB 10; Length 406;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
Db      45 TAAAAAAAAAAAAAAAAAACTCGAG 20
```

```
RESULT 24
US-09-918-995-19475
; Sequence 19475, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19475
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(448)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-19475
```

```
Query Match          2.0%; Score 26; DB 11; Length 448;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
Db      364 TAAAAAAAAAAAAAAAAAACTCGAG 389
```

```
RESULT 25
US-09-918-995-22561
; Sequence 22561, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 22561
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(475)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-22561
```

```
Query Match          2.0%; Score 26; DB 11; Length 475;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
Db      427 TAAAAAAAAAAAAAAAAAACTCGAG 452
```

```
RESULT 26
US-09-918-995-23066
; Sequence 23066, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23066
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(502)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23066
```

```
Query Match          2.0%; Score 26; DB 11; Length 502;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
Db      463 TAAAAAAAAAAAAAAAAAACTCGAG 488
```

```
RESULT 27
US-09-918-995-19781
; Sequence 19781, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19781
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

NAME/KEY: misc_feature
LOCATION: (1)...(513)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-19781

Query Match 2.0%; Score 26; DB 11; Length 513;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 485 TAAAAAAAAAAAAAAAACTCGAG 510

RESULT 28
US-09-764-887-134/C
Sequence 134, Application US/09764887
Patent No. US20020042096A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P113
CURRENT APPLICATION NUMBER: US/09/764,887
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 658
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 134
LENGTH: 519
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (506)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-887-134

Query Match 2.0%; Score 26; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 44 TAAAAAAAAAAAAAAAACTCGAG 19

RESULT 29
US-10-073-961-134/C
Sequence 134, Application US/10073961
Publication No. US2003007602A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P113C1
CURRENT APPLICATION NUMBER: US/10/073,961
CURRENT FILING DATE: 2002-03-05
Prior application data removed - consult PALM or file wrapper
Prior Application Number: 09/764,887
Prior Filing Date: 2001-01-17
Prior Application Number: 60/179,065
Prior Filing Date: 2000-01-31
Prior Application Number: 60/180,628
Prior Filing Date: 2000-02-04
Prior Application Number: 60/214,886
Prior Filing Date: 2000-06-28
Prior Application Number: 60/217,487
Prior Filing Date: 2000-07-11
Prior Application Number: 60/225,758
Prior Filing Date: 2000-08-14
Prior Application Number: 60/220,963
Prior Filing Date: 2000-07-26
Prior Application Number: 60/217,496
Prior Filing Date: 2000-07-11
Prior Application Number: 60/225,447

Prior Filing Date: 2000-08-14
Prior Application Number: 60/218,290
Prior Filing Date: 2000-07-14
Prior Application Number: 60/225,757
Prior Filing Date: 2000-08-14
Prior Application Number: 60/226,868
Prior Filing Date: 2000-08-22
Prior Application Number: 60/216,647
Prior Filing Date: 2000-07-07
Prior Application Number: 60/225,267
Prior Filing Date: 2000-08-14
Prior Application Number: 60/216,880
Prior Filing Date: 2000-07-07
Prior Application Number: 60/225,270
Prior Filing Date: 2000-08-14
Prior Application Number: 60/251,869
Prior Filing Date: 2000-12-08
Prior Application Number: 60/235,834
Prior Filing Date: 2000-09-27
Prior Application Number: 60/234,274
Prior Filing Date: 2000-09-21
Prior Application Number: 60/234,223
Prior Filing Date: 2000-09-21
Prior Application Number: 60/228,924
Prior Filing Date: 2000-08-30
Prior Application Number: 60/224,518
Prior Filing Date: 2000-08-14
Prior Application Number: 60/236,369
Prior Filing Date: 2000-09-29
Prior Application Number: 60/224,519
Prior Filing Date: 2000-08-14
Prior Application Number: 60/220,964
Prior Filing Date: 2000-07-26
Prior Application Number: 60/241,809
Prior Filing Date: 2000-10-20
Prior Application Number: 60/249,299
Prior Filing Date: 2000-11-17
Prior Application Number: 60/236,327
Prior Filing Date: 2000-09-29
Prior Application Number: 60/241,785
Prior Filing Date: 2000-10-20
Prior Application Number: 60/244,617
Prior Filing Date: 2000-11-01
Prior Application Number: 60/225,268
Prior Filing Date: 2000-08-14
Prior Application Number: 60/236,368
Prior Filing Date: 2000-09-29
Prior Application Number: 60/251,856
Prior Filing Date: 2000-12-08
Prior Application Number: 60/251,868
Prior Filing Date: 2000-12-08
Prior Application Number: 60/229,344
Prior Filing Date: 2000-09-01
Prior Application Number: 60/234,997
Prior Filing Date: 2000-09-25
Prior Application Number: 60/229,343
Prior Filing Date: 2000-09-01
Prior Application Number: 60/229,345
Prior Filing Date: 2000-09-01
Prior Application Number: 60/229,287
Prior Filing Date: 2000-09-01
Prior Application Number: 60/229,513
Prior Filing Date: 2000-09-05
Prior Application Number: 60/231,413
Prior Filing Date: 2000-09-08
Prior Application Number: 60/229,509
Prior Filing Date: 2000-09-05
Prior Application Number: 60/236,367
Prior Filing Date: 2000-09-29
Prior Application Number: 60/237,039
Prior Filing Date: 2000-10-02
Prior Application Number: 60/237,038
Prior Filing Date: 2000-10-02

PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,081

PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08

Query Match 2.0%; Score 26; DB 15; Length 519;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTGAG 1323
Db 44 TAAAAAAAAAAAAAAAACTGAG 19

RESULT 30
US-10-106-698-1187
; Sequence 1187, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: P8005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1187
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (587)..(587)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1187

Query Match 2.0%; Score 26; DB 15; Length 591;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1298 TAAAAAAAAAAAAAAAACTGAG 1323
|||||

Db 545 TAAAAAAAAAAAAAAAAAATCGAG 570

RESULT 31

US-10-106-698-517
; Sequence 517, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 517
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (619)..(619)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-517

Query Match

Best Local Similarity 100.0%; Score 26; DB 15; Length 646;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
Db 569 TAAAAAAAAAAAAAAAAAATCGAG 594

RESULT 32

US-09-832-129-16
; Sequence 16, Application US/09832129
; Publication No. US20030027297A1
; GENERAL INFORMATION:
; APPLICANT: Fliscella et al.
; TITLE OF INVENTION: 19 Human secreted proteins
; FILE REFERENCE: P2045P1
; CURRENT APPLICATION NUMBER: US/09/832,129
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: PCT/US00/28664
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/163,085
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 60/172,411
; PRIOR FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-129-16

Query Match

Best Local Similarity 100.0%; Score 26; DB 11; Length 682;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
Db 644 TAAAAAAAAAAAAAAAAAATCGAG 669

RESULT 33

US-10-342-224-3
; Sequence 3, Application US/10342224
; Publication No. US20030162294A1
; GENERAL INFORMATION:
; APPLICANT: Nathalie Verbruggen
; TITLE OF INVENTION: Genes Involved in Tolerance to Environmental Stress
; FILE REFERENCE: CN-012US
; CURRENT APPLICATION NUMBER: US/10/342,224
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US/09/762,154
; PRIOR FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: EP 98202634.6
; PRIOR FILING DATE: 1998-08-04
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 695
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)...(564)
US-10-342-224-3

Query Match

Best Local Similarity 100.0%; Score 26; DB 13; Length 695;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1298 TAAAAAAAAAAAAAAAAAATCGAG 1323
Db 669 TAAAAAAAAAAAAAAAAAATCGAG 694

RESULT 34

US-09-966-881-26/c
; Sequence 26, Application US/09966881
; Patent No. US20020120960A1
; GENERAL INFORMATION:
; APPLICANT: Seymour, Graham
; Bird, Colin
; Medina-Suarez, Rosybel
; TITLE OF INVENTION: Genetic control of Fruit Ripening
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zeneca Ag Products Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/966,881
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,860
; FILING DATE: 29-Mar-1999
; APPLICATION NUMBER: GB 9618862.8
; FILING DATE: 10-Sep-1996
; APPLICATION NUMBER: GB 9708366.1
; FILING DATE: 25-Apr-1997
; APPLICATION NUMBER: PCT/GB97/02424
; FILING DATE: 08-Sep-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoheneshtz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: SEE 50183/UST


```
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (382) 886-1699
  INFORMATION FOR SEQ ID NO: 26:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 786 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: unknown
  MOLECULE TYPE: cDNA
  IMMEDIATE SOURCE:
    CLONE: U-0125
  SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-881-26

Query Match
Best Local Similarity 100.0%; Score 26; DB 10; Length 786;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAATCTCGAG 1323
Db 39 TAAAAAAAAAAAAAAAAAATCTCGAG 14

RESULT 35
US-10-027-632-155629
  Sequence 155629, Application US/10027632
  Publication No. US20030204075A9
  GENERAL INFORMATION:
    APPLICANT: Wang, David G.
    TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
    FILE REFERENCE: 108827.129
    CURRENT APPLICATION NUMBER: US/10/027,632
    CURRENT FILING DATE: 2002-04-30
    PRIOR APPLICATION NUMBER: US 60/218,006
    PRIOR FILING DATE: 2000-07-12
    PRIOR APPLICATION NUMBER: US 60/198,676
    PRIOR FILING DATE: 2000-04-20
    PRIOR APPLICATION NUMBER: US 60/193,483
    PRIOR FILING DATE: 2000-03-29
    PRIOR APPLICATION NUMBER: US 60/185,218
    PRIOR FILING DATE: 2000-02-24
    PRIOR APPLICATION NUMBER: US 60/167,363
    PRIOR FILING DATE: 1999-11-23
    PRIOR APPLICATION NUMBER: US 60/156,358
    PRIOR FILING DATE: 1999-09-28
    PRIOR APPLICATION NUMBER: US 60/146,002
    PRIOR FILING DATE: 1999-08-09
    NUMBER OF SEQ ID NOS: 325720
    SOFTWARE: FastSeq for Windows Version 4.0
    SEQ ID NO 155629
    LENGTH: 825
    TYPE: DNA
    ORGANISM: Human
  US-10-027-632-155629

Query Match
Best Local Similarity 100.0%; Score 26; DB 13; Length 825;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 155629
LENGTH: 825
TYPE: DNA
ORGANISM: Human
US-10-027-632-155629

Query Match
Best Local Similarity 100.0%; Score 26; DB 14; Length 825;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1292 CTCAGTAAAAAAAAAAAAAAAAA 1317
Db 77 CTCAGTAAAAAAAAAAAAAAAAA 102

RESULT 37
US-10-264-237-135
  Sequence 135, Application US/10264237
  Publication No. US2004009491A1
  GENERAL INFORMATION:
    APPLICANT: Birse et al.
    TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
    FILE REFERENCE: PA131PI
    CURRENT APPLICATION NUMBER: US/10/264,237
    CURRENT FILING DATE: 2002-10-04
    PRIOR APPLICATION NUMBER: PCT/US01/16450
    PRIOR FILING DATE: 2001-05-18
    PRIOR APPLICATION NUMBER: US 60/205,515
    PRIOR FILING DATE: 2000-05-19
    NUMBER OF SEQ ID NOS: 2876
    SOFTWARE: PatentIn Ver. 3.1
    SEQ ID NO 135
    LENGTH: 1076
    TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
      NAME/KEY: misc.feature
      LOCATION: (35)..(35)
    OTHER INFORMATION: n equals a,t,g, or c
  US-10-264-237-135

Query Match
Best Local Similarity 100.0%; Score 26; DB 12; Length 1076;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAATCTCGAG 1323
Db 1051 TAAAAAAAAAAAAAAAAAATCTCGAG 1076

RESULT 38
US-09-872-523-4
  Sequence 4, Application US/09872523
  Patent No. US20020137906A1
  GENERAL INFORMATION:
```

```

; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1108
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-4

Query Match      2.0%; Score 26; DB 10; Length 1108;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
Db      1083 TAAAAAAAAAAAAAAAAAAGCTCGAG 1108

RESULT 39
US-09-872-523-48
; Sequence 48, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 1108
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-48

Query Match      2.0%; Score 26; DB 10; Length 1108;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
Db      1083 TAAAAAAAAAAAAAAAAAAGCTCGAG 1108

RESULT 40
US-09-814-353-21392
; Sequence 21392, Application US/09814353
; Publication No. US2003015831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Little, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
```

```

; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21392
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 249, 1102, 1103, 1104, 1105, 1106, 1107, 1108, 1109, 1110,
; LOCATION: 1111, 1112, 1113, 1114, 1115, 1116, 1117, 1118, 1119, 1120,
; LOCATION: 1121, 1122, 1123, 1124, 1125, 1126, 1127, 1128, 1129, 1130,
; LOCATION: 1131, 1132, 1133, 1134, 1135, 1136, 1137, 1138, 1139
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1140, 1141, 1142, 1143
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21392

Query Match      2.0%; Score 26; DB 13; Length 1143;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
Db      934 TAAAAAAAAAAAAAAAAAAGCTCGAG 959

RESULT 41
US-09-738-973-27
; Sequence 27, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Flinn, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrila, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 1331
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-973-27

Query Match      2.0%; Score 26; DB 10; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1298 TAAAAAAAAAAAAAAAAAAGCTCGAG 1323
```

Db 1306 TAAAAAAAAAAAAAAAACTCGAG 1331

RESULT 42

US-09-854-133-27
; Sequence 27, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 1331
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-27

Query Match 2.0%; Score 26; DB 10; Length 1331;

Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1306 TAAAAAAAAAAAAAAAACTCGAG 1331

RESULT 43

US-10-144-649A-27
; Sequence 27, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 1331
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-649A-27

Query Match 2.0%; Score 26; DB 15; Length 1331;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1306 TAAAAAAAAAAAAAAAACTCGAG 1331

RESULT 44

US-10-231-417-129
; Sequence 129, Application US/10231417
; Publication No. US20030176681A1
; GENERAL INFORMATION:

APPLICANT: Feng et al.
; TITLE OF INVENTION: 148 Human Secreted Proteins
; FILE REFERENCE: P2019P1
; CURRENT APPLICATION NUMBER: US/10/231,417
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US/09/296,622
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 619
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
; LENGTH: 1334
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-231-417-129

Query Match 2.0%; Score 26; DB 13; Length 1334;

Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1309 TAAAAAAAAAAAAAAAACTCGAG 1334

RESULT 45

US-09-925-299-82
; Sequence 82, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-299-82

Query Match 2.0%; Score 26; DB 9; Length 1350;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAAAAAAAAAAAACTCGAG 1323
Db 1324 TAAAAAAAAAAAAAAAACTCGAG 1349

RESULT 46

US-09-925-299-82
; Sequence 82, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 1350

TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-299-82

Query Match 2.0%; Score 26; DB 11; Length 1350;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAACTCGAG 1323
Db 1324 TAAAAAACTCGAG 1349

RESULT 47
US-10-106-698-2096

Sequence 2096, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698

CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280

PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0

SEQ ID NO 2096
LENGTH: 1350
TYPE: DNA

ORGANISM: Homo sapiens
US-10-106-698-2096

Query Match 2.0%; Score 26; DB 15; Length 1350;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAACTCGAG 1323
Db 1324 TAAAAAACTCGAG 1349

RESULT 48
US-09-809-391-207

Sequence 207, Application US/09809391
Publication No. US20030049618A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P2
CURRENT APPLICATION NUMBER: US/09/809,391

CURRENT FILING DATE: 2001-03-16
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 761
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 207
LENGTH: 1354
TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: SITE
LOCATION: (465)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (794)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (1344)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (1349)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (1350)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (1352)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (1354)
OTHER INFORMATION: n equals a,t,g, or c

US-09-809-391-207

Query Match 2.0%; Score 26; DB 11; Length 1354;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1298 TAAAAAACTCGAG 1323
Db 1287 TAAAAAACTCGAG 1312

RESULT 49
US-09-882-171-207

Sequence 207, Application US/09882171
Publication No. US20030175858A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P2
CURRENT APPLICATION NUMBER: US/09/882,171

CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 09/809,391
PRIOR FILING DATE: 2001-03-16

PRIOR APPLICATION NUMBER: 09/149,476
PRIOR FILING DATE: 1998-09-08
PRIOR APPLICATION NUMBER: PCT/US98/04493

PRIOR FILING DATE: 1998-03-06
PRIOR APPLICATION NUMBER: 60/040,162
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/040,333
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/038,621

PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/040,626
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/040,334
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/040,336

PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/040,163
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/047,600
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,615

PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,597
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,502
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,633

PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,583
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,617
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,618

PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,503
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,592
PRIOR FILING DATE: 1997-05-23

; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/057,669
; PRIOR FILING DATE: 1997-09-05

Query Match 2.0%; Score 26; DB 13; Length 1354;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 1287 TAAAAAAAAAAAAAAAAAACTCGAG 1312

RESULT 50
US-10-050-704-92
; Sequence 92, Application US/10050704
; Publication No. US2003050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 92
; LENGTH: 1433
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-050-704-92

Query Match 2.0%; Score 26; DB 15; Length 1433;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1298 TAAAAAAAAAAAAAAAAAACTCGAG 1323
DB 1392 TAAAAAAAAAAAAAAAAAACTCGAG 1417

Search completed: January 28, 2004, 16:03:45
Job time : 446 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 15:56:05 ; Search time 2681 Seconds
(without alignments)
11993.593 Million cell updates/sec

Title: US-09-881-556A-3
Perfect score: 1323
Sequence: 1 gaattcgagcagattttt.....aaaaaaaaaacctcgag 1323

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 15
Total number of hits satisfying chosen parameters: 2201226

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

EST.*
1: em_gatba:*
2: em_gatbun:*
3: em_gatln:*
4: em_gatnu:*
5: em_gatov:*
6: em_gatp1:*
7: em_gatp2:*
8: em_gatp3:*
9: gb_gat1:*
10: gb_gat2:*
11: gb_gat3:*
12: gb_gat4:*
13: gb_gat5:*
14: gb_gat6:*
15: em_gatfun:*
16: em_gatcom:*
17: em_gat_hum:*
18: em_gat_hiv:*
19: em_gat_pln:*
20: em_gat_vrt:*
21: em_gat_fun:*
22: em_gat_man:*
23: em_gat_mus:*
24: em_gat_pro:*
25: em_gat_rtd:*
26: em_gat_phg:*
27: em_gat_vrl:*
28: gb_gat1:*
29: gb_gat2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	193	14.6	837 28	AQ411949
2	179	13.5	580 29	CPA562092
3	32	2.4	683 12	BM065008
4	31	2.3	300 9	AJ484699

C 5	31	2.3	342	9	AW218331	AW218331 EST303514
C 6	31	2.3	360	9	AJ484700	AJ484700 AJ484700
C 7	31	2.3	686	13	BU060709	BU060709 Fgr-C-1 H
C 8	31	2.3	689	13	BU061154	BU061154 Fgrn-5 E1
C 9	31	2.3	846	13	BO144248	BO144248 NF055G1D
C 10	30	2.3	407	14	CA969613	CA969613 wlmk8.pko
C 11	30	2.3	447	9	AA126428	AA126428 zn84h09.s
C 12	30	2.3	515	13	BU721218	BU721218 SJMB2KH09
C 13	30	2.3	559	9	BO081761	BO081761 san28h07.
C 14	30	2.3	603	9	AW334649	AW334649 SJ8G4 AGS
C 15	30	2.3	647	10	BE661929	BE661929 1240 Gmax
C 16	30	2.3	678	10	BE641651	BE641651 NF051F01
C 17	30	2.3	867	10	BF316408	BF316408 601901981
C 18	29	2.2	166	13	BM130194	BM130194 BM130194
C 19	29	2.2	259	14	CA671233	CA671233 wlsu2.pko
C 20	29	2.2	261	14	CA729647	CA729647 wlp1c.pko
C 21	29	2.2	269	14	BE232362	BE232362 paf0a04.y
C 22	29	2.2	294	9	AA325025	AA325025 EST27949
C 23	29	2.2	301	12	BI673873	BI673873 t330b12.y
C 24	29	2.2	332	12	BM530269	BM530269 fw74d10.y
C 25	29	2.2	364	12	BI596045	BI596045 949076802
C 26	29	2.2	367	12	BI233332	BI233332 949010A01
C 27	29	2.2	377	10	BE445889	BE445889 WHE1148.E
C 28	29	2.2	407	12	BI674867	BI674867 949076802
C 29	29	2.2	438	14	CA403603	CA403603 EL01N0504
C 30	29	2.2	472	12	BP123504	BP123504 BP123504
C 31	29	2.2	487	10	AW967310	AW967310 EST379385
C 32	29	2.2	490	12	BI165112	BI165112 saf04c06.
C 33	29	2.2	516	12	BP122779	BP122779 BP122779
C 34	29	2.2	516	12	BP122909	BP122909 BP122909
C 35	29	2.2	516	12	BP123880	BP123880 BP123880
C 36	29	2.2	516	12	BP123966	BP123966 BP123966
C 37	29	2.2	516	12	BP124078	BP124078 BP124078
C 38	29	2.2	516	12	BP124196	BP124196 BP124196
C 39	29	2.2	516	12	BP124420	BP124420 BP124420
C 40	29	2.2	518	12	BM303636	BM303636 SNEST4a62
C 41	29	2.2	519	9	A1880998	A1880998 T7766 MVA
C 42	29	2.2	524	9	AV699251	AV699251 AV699251
C 43	29	2.2	525	12	BP126517	BP126517 BP126517
C 44	29	2.2	528	12	BP123680	BP123680 BP123680
C 45	29	2.2	544	12	BP123416	BP123416 BP123416
C 46	29	2.2	577	9	A1295021	A1295021 LP08531.5
C 47	29	2.2	588	10	AW953985	AW953985 EST365950
C 48	29	2.2	600	13	BU922191	BU922191 7021-08 M
C 49	29	2.2	604	9	A1960332	A1960332 sc81h08.y
C 50	29	2.2	617	10	AW949436	AW949436 EST361506
C 51	29	2.2	622	14	CB086597	CB086597 hj87a11.9
C 52	29	2.2	630	12	BG339807	BG339807 cr62f05.x
C 53	29	2.2	656	9	AW331996	AW331996 S1F1 AGS-
C 54	29	2.2	667	10	BG454533	BG454533 NF112B10L
C 55	29	2.2	685	10	BG454873	BG454873 NF108H05L
C 56	29	2.2	704	13	BO797243	BO797243 EST_6181
C 57	29	2.2	783	13	BO151716	BO151716 NF022H01I
C 58	29	2.2	845	12	BP128084	BP128084 BP128084
C 59	29	2.2	856	10	BF312670	BF312670 601898169
C 60	29	2.2	921	13	BO677312	BO677312 AGENCOURT
C 61	29	2.2	927	13	BU791426	BU791426 SJFAYB02
C 62	29	2.2	967	12	BO061895	BO061895 AGENCOURT
C 63	29	2.2	1106	12	BM911170	BM911170 AGENCOURT
C 64	28	2.1	147	12	BM571016	BM571016 fx04a02.y
C 65	28	2.1	198	12	BM072073	BM072073 fu88a03.y
C 66	28	2.1	217	9	AW410991	AW410991 fh09f08.y
C 67	28	2.1	252	14	CD478266	CD478266 gca01-40m
C 68	28	2.1	253	14	CA519203	CA519203 KS1002280
C 69	28	2.1	270	14	CD405844	CD405844 Gm_CK292
C 70	28	2.1	273	9	AL374286	AL374286 MCBBO5G02
C 71	28	2.1	295	13	BO152853	BO152853 NF025H03I
C 72	28	2.1	314	14	CA697780	CA697780 w1k4.pko0
C 73	28	2.1	344	9	A1781084	A1781084 EST261963
C 74	28	2.1	368	9	AL382543	AL382543 MCBCO7H09
C 75	28	2.1	372	10	BE352543	BE352543 pa34f11.y
C 76	28	2.1	375	12	BM252771	BM252771 SNEST4a27
C 77	28	2.1	375	14	CA654905	CA654905 w1m0.pko0

C 78	28	2.1	384	10	BE050094	BE050094 za56f01.g
C 79	28	2.1	386	9	AV566378	AV566378 AV566378
C 80	28	2.1	393	9	AL377177	AL377177 MEBB29H02
C 81	28	2.1	395	10	BF635108	BF635108 NF07G002D
C 82	28	2.1	397	13	BQ796462	BQ796462 EST 5400
C 83	28	2.1	402	9	AV556169	AV556169 AV556169
C 84	28	2.1	403	10	BF716479	BF716479 NCES3a13
C 85	28	2.1	414	12	BI432719	BI432719 EST535480
C 86	28	2.1	425	14	CA520651	CA520651 K811015C0
C 87	28	2.1	428	14	CA663540	CA663540 wimk1.pk0
C 88	28	2.1	432	9	AL378662	AL378662 MEBB39F10
C 89	28	2.1	434	13	BQ799597	BQ799597 EST 1766
C 90	28	2.1	442	9	AM023070	AM023070 d148e06.y
C 91	28	2.1	448	10	BE420260	BE420260 WMS04.E39R
C 92	28	2.1	456	9	AL385465	AL385465 MEBB28G03
C 93	28	2.1	459	13	BU495749	BU495749 PEST0ab7
C 94	28	2.1	461	12	BM072010	BM072010 f187a05.y
C 95	28	2.1	461	14	CA668386	CA668386 w1su1.pk0
C 96	28	2.1	469	9	AL384638	AL384638 MEBB23E02
C 97	28	2.1	474	10	BE050715	BE050715 za68g02.b
C 98	28	2.1	480	10	BG356100	BG356100 EM1 21.D0
C 99	28	2.1	488	9	AL371534	AL371534 MEBB44H01
C 100	28	2.1	490	9	AM699845	AM699845 gb33e11.y

ALIGNMENTS

RESULT 1
 AQ411949
 LOCUS
 DEFINITION Cg0925A CpiOwAGDNa1 Cryptosporidium parvum genomic, genomic survey sequence.
 AQ411949
 ACCESSION AQ411949.1 GI:4469673
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CRYPTOSPORIDIUM PARVUM
 Cryptosporidium parvum
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Cryptosporidiidae; Cryptosporidium.
 1 (bases 1 to 837)
 Strong, W.B. and Nelson, R.G.
 Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis
 Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 10717299
 Contact: Nelson, R. G.
 Dept. of Medicine & Pharmaceutical Chemistry
 San Francisco General Hospital-University of California, San Francisco
 Box 0811, San Francisco, CA 94143-0811, USA
 Tel: 415 206 8846
 Fax: 415 206 3353
 Email: malari@itsa.ucsf.edu
 For Annotation Data see http://medefgh.ucsf.edu/id/CpTags/home.html
 Seq primer: M13(-21) forward
 Class: shotgun

FEATURES

source
 1. 837
 /organism="Cryptosporidium parvum"
 /mol_type="genomic DNA"
 /strain="IOWA"
 /db_xref="taxon:5807"
 /lab_host="E. coli XL2 Blue MRF"
 /clone_id="CpiOwAGDNa1"
 /note="Vector: Bluescript II (SK-); Site 1: EcoRV, C. parvum (IOWA isolate) genomic DNA was hydrolonomically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center (http://sequence-www.stanford.edu/group/techdev/shear.htm). The randomly sheared gDNA was chromatographed on

Seaphacryl S-400 to remove any small fragments and DNA eluting in the void volume was subcloned into an EcoR V-digested, alkaline phosphatase-treated pluscript II (SK-) vector and transformed into E. coli strain XL2 Blue MRF. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

BASE COUNT 333 a 128 c 124 g 252 t

Query Match 14.6%; Score 193; DB 28; Length 837;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 293; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1008	AAATCAAGCCAGCTACATATCTTCAACCAAGAAATGAATCATATATACGAGAGG	1067
DB	119	AAATCAAGCCAGCTACATATCTTCAACCAAGAAATGAATCATATATACGAGAGG	178
QY	1068	TAATTCAGTTTAAATPACTAAATCTTCCAAATCAACCCATCTTTCACGCAAGCTGGC	1127
DB	179	TAATTCAGTTTAAATPACTAAATCTTCCAAATCAACCCATCTTTCACGCAAGCTGGC	238
QY	1128	AGATGTGGAAGTTATGAGGAAAAAGATAAATGCTGTGAGTGAGAGTAATCACTAA	1187
DB	239	AGATGTGGAAGTTATGAGGAAAAAGATAAATGCTGTGAGTAATCACTAA	298
QY	1188	AAATGACATTTATGAGATCTCTGAAAGAAATTAAGAAATTAAGATATCTGCAT	1247
DB	299	AAATGACATTTATGAGATCTCTGAAAGAAATTAAGAAATTAAGATATCTGCAT	358
QY	1248	TGATTAATATGAGCTTATGAGCTTATGAGCTTATGAGCTTATGAGCTTATGAGCTT	1302
DB	359	TGATTAATATGAGCTTATGAGCTTATGAGCTTATGAGCTTATGAGCTTATGAGCTT	413

RESULT 2
 CPAS62092
 LOCUS
 DEFINITION CPAS62092 580 bp DNA linear GSS 02-JUN-2003
 sequence, genomic survey sequence.
 CPAS62092
 ACCESSION AF562092.1 GI:31336694
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CRYPTOSPORIDIUM PARVUM
 Cryptosporidium parvum
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Cryptosporidiidae; Cryptosporidium.
 1
 Bankier, A.T., Spriggs, H.F., Partmann, B., Konfortov, B.A., Madera, M., Vogel, C., Teichmann, S.A., Ivens, A. and Dear, P.H.
 Integrated mapping, chromosomal sequencing and sequence analysis of Cryptosporidium parvum
 Genome Res. (2003) In press
 2 (bases 1 to 580)
 Dear, P.H.
 Direct Submission
 Submitted (09-MAY-2003) Dear P.H., PNAC Biotech Division, MRC Laboratory of Molecular Biology, Hills Road, Cambridge, Cambs CB2 2QH, UNITED KINGDOM

FEATURES

source
 1. 580
 /organism="Cryptosporidium parvum"
 /mol_type="genomic DNA"
 /serotype="Type 2"
 /isolate="IOWA"
 /db_xref="taxon:5807"
 /chromosome="7"
 /clone="pica_0006_901"
 /clone_id="pica"
 /dev_stage="oocyst"
 /note="7T end sequence"
 BASE COUNT 231 a 87 c 91 g 170 t 1 others
 ORIGIN

Query Match 13.5%; Score 179; DB 29; Length 580;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1124 TGGCAGATGTGGAAGATTATGGGAAAGATAAAATGTTGATGTGAGCAAGTATCA 1183
DB 187 TGGCAGATGTGGAAGATTATGGGAAAGATAAAATGTTGATGTGAGCAAGTATCA 246
QY 1184 CTAAATATGACATATTATGAGATCTCGAAGAAATTAGAAACAAATGTAAATATCT 1243
DB 247 CTAAATATGACATATTATGAGATCTCGAAGAAATTAGAAACAAATGTAAATATCT 306
QY 1244 GCATTGATAAATATGCGCTTACCATTTCCAAATATCTAAATGTCACTCAAGTAAA 1302
DB 307 GCATTGATAAATATGCGCTTACCATTTCCAAATATCTAAATGTCACTCAAGTAAA 365

RESULT 3
BM065008/c 683 bp mRNA linear EST 11-SEP-2002
LOCUS KS0107306 KS01 Capsicum annuum cDNA, mRNA sequence.
DEFINITION BM065008
ACCESSION BM065008
VERSION BM065008.1 GI:22785126
KEYWORDS EST.
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Lamiales; Solanales; Solanaceae; Capsicum.
1 (bases 1 to 683)
Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pal, H.-S., Hur, C.-G. and Choi, D.
Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen
Unpublished
CONTACT: Doil Choi
Genome Research Center and National Center for Genome Information
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.krdb.re.kr
High quality sequence stop: 683.
Location/Qualifiers
1..683
/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Bukang"
/db_xref="taxon:4072"
/cvs_type="leaf inoculated with Xanthomonas campestris pv. glycines"
/dev_stage="8 weeks after germination"
/clone_id="KS01"
/note="Vector: pBluescript SK(-)"

BASE COUNT 208 a 146 c 83 g 246 t

ORIGIN

Query Match 2.4%; Score 32; DB 12; Length 683;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1292 CTCAGTAAAAAATCTCGAG 1323
DB 32 CTCAGTAAAAAATCTCGAG 1

RESULT 4
AJ484699 300 bp mRNA linear EST 24-MAY-2002
LOCUS AJ484699 S00011 Hordeum vulgare cDNA clone S000110029F03F1, mRNA sequence.

ACCESSION AJ484699
VERSION AJ484699.1 GI:2120655
KEYWORDS EST.
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
1 (bases 1 to 300)
Saren, A.-M., Tanekenen, J., Paulin, L. and Schulman, A.H.
Barley EST's
Unpublished
CONTACT: Schulman AH
Institute of Biotechnology
University of Helsinki
P.O. Box 56 (Valkankari 6A), University of Helsinki FIN-00014, Finland.
Location/Qualifiers
1..300
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
/clone="S000110029F03F1"
/dev_stage="Developing seed"
/clone_id="S00011"
/note="12,15,18 days after pollination"

BASE COUNT 87 a 66 c 54 g 93 t

ORIGIN

Query Match 2.3%; Score 31; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1293 TCAGTAAAAAATCTCGAG 1323
DB 42 TCAGTAAAAAATCTCGAG 12

RESULT 5
AM218331/c 342 bp mRNA linear EST 18-MAY-2001
LOCUS EST303514 tomato radicle, 5 d post-imbibition, Cornell University
DEFINITION Lycopersicon esculentum cDNA clone cLEZ6021 similar to extensin-like protein, Df10, mRNA sequence.
ACCESSION AM218331
VERSION AM218331.1 GI:6529205
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato).
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 342)
van der Hoeven, R.S., Garvin, D., Matern, A.L., Holt, I.E., Liang, F., Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Romling, C.M., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
Generation of ESTs from tomato radicle tissue
Unpublished
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
3 prime sequence.
Location/Qualifiers
1..342
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEZ6021"
/cvs_type="radicle"
/dev_stage="Seedlings 5 days post-imbibition"

/clone_1lb="tomato radicle, 5 d post-imbibition, Cornell University"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Tanskley; Tissue supplied by Dave Garvin (USDA-ARS, Ithaca, NY 14850)."

BASE COUNT 137 a 36 c 70 g 99 t

ORIGIN

Query Match 2.3%; Score 31; DB 9; Length 342;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1293 TCAAGTAAAAAAAAAAAAAAAACTCGAG 1323
 61 TCAAGTAAAAAAAAAAAAAAAACTCGAG 31

Db

RESULT 6
 AJ484700/c 360 bp mRNA linear EST 24-MAY-2002
 LOCUS AJ484700 S00011 Hordeum vulgare CDNA clone S0001100038A06F1, mRNA
 DEFINITION
 accession AJ484700.1 GI:21200656
 VERSION AJ484700.1 GI:21200656
 KEYWORDS EST.
 SOURCE Hordeum vulgare
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 /triticaceae; Hordeum.
 1 (bases 1 to 360)
 Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
 Barley EST's
 TITLE Unpublished
 JOURNAL
 COMMENT Contact: Schulman AH
 Institute of Biotechnology
 University of Helsinki
 P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014, Finland.

FEATURES
 source
 1..360
 /organism="Hordeum vulgare"
 /mol_type="mRNA"
 /db_xref="taxon:4513"
 /clone="S0001100038A06F1"
 /dev_string="developing seed"
 /clone_1lb="S00011"
 /note="12,15,18 days after pollination"
 83 c 67 g 101 t

BASE COUNT 109 a 83 c 67 g 101 t

ORIGIN

Query Match 2.3%; Score 31; DB 9; Length 360;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1293 TCAAGTAAAAAAAAAAAAAAAACTCGAG 1323
 42 TCAAGTAAAAAAAAAAAAAAAACTCGAG 12

Db

RESULT 7
 BU060709/c 686 bp mRNA linear EST 26-AUG-2002
 LOCUS BU060709
 DEFINITION Fgr-C_1_H23_T7 Carbon-starved mycelia Gibberella zeae cDNA, mRNA
 accession BU060709.1 GI:22500998
 VERSION BU060709.1 GI:22500998
 KEYWORDS EST.
 SOURCE Gibberella zeae
 ORGANISM Gibberella zeae
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 1 (bases 1 to 686)

REFERENCE 1 (bases 1 to 686)

AUTHORS Trail,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C.
 TITLE Analysis of expressed sequence tags from Gibberella zeae (anamorph
 Fusarium graminearum)
 JOURNAL Fungal Genet. Biol. 38 (2), 187-197 (2003)
 MEDLINE 22508120
 PUBMED 12620255

COMMENT Contact: Frances Trail
 Department of Plant Biology
 Michigan State University
 East Lansing, MI 48824, USA
 Tel: 517 432 2939
 Fax: 517 353 1926
 Email: trail@msu.edu
 Plate: 1 row: H column: 23.
 Location/Qualifiers

FEATURES
 source
 1..686
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="NRRL 31084"
 /db_xref="taxon:5518"
 /clone_1lb="Nitrogen-starved mycelia"
 /note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 151 a 180 c 186 g 172 t

ORIGIN

Query Match 2.3%; Score 31; DB 13; Length 689;

RESULT 8
 BU061154/c 689 bp mRNA linear EST 26-AUG-2002
 LOCUS BU061154
 DEFINITION FgrN 5_E15_T7 Nitrogen-starved mycelia Gibberella zeae cDNA, mRNA
 accession BU061154.1 GI:22501443
 VERSION BU061154.1 GI:22501443
 KEYWORDS EST.
 SOURCE Gibberella zeae
 ORGANISM Gibberella zeae
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 1 (bases 1 to 689)
 Trail,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C.
 Analysis of expressed sequence tags from Gibberella zeae (anamorph
 Fusarium graminearum)
 JOURNAL Fungal Genet. Biol. 38 (2), 187-197 (2003)
 MEDLINE 22508120
 PUBMED 12620255

COMMENT Contact: Frances Trail
 Department of Plant Biology
 Michigan State University
 East Lansing, MI 48824, USA
 Tel: 517 432 2939
 Fax: 517 353 1926
 Email: trail@msu.edu
 Plate: 5 row: E column: 15.
 Location/Qualifiers

FEATURES
 source
 1..689
 /organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="NRRL 31084"
 /db_xref="taxon:5518"
 /clone_1lb="Nitrogen-starved mycelia"
 /note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 151 a 180 c 186 g 172 t

ORIGIN

Query Match 2.3%; Score 31; DB 13; Length 689;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1293 TCAGTAAAAAATCTCGAG 1323
|||||
40 TCAGTAAAAAATCTCGAG 10

RESULT 9
LOCUS B0144248 846 bp mRNA linear EST 24-APR-2002
DEFINITION NF05G11DT.F1088 Drought Medicago truncatula cDNA clone NF05G11DT
5', mRNA sequence.
B0144248
B0144248.1 GI:20281307
EST.
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago
1 (bases 1 to 846)
Torres-Varela, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Iman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula drought library
Unpublished
JOURNAL Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert Length: 846 Std Error: 0.00
Plate: 055 Row: G Column: 11
Seq primer: TCACACGAGAAACAGCTATGAC.
location/Qualifiers
FEATURES
source
1..846
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF05G11DT"
/cisue_type="plantlet"
/dev_stage="pooled timepoints"
/clone_lib="Drought"
/note="Vector: Lambda Zap; Contains a mixture of entire
plantlets harvested in a series of days-post-watering
timepoints."
BASE COUNT 231 a 204 c 181 g 230 t
ORIGIN

Query Match 2.3%; Score 31; DB 13; Length 846;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1293 TCAGTAAAAAATCTCGAG 1323
|||||
607 TCAGTAAAAAATCTCGAG 637

RESULT 10
LOCUS CA696313/c 407 bp mRNA linear EST 26-NOV-2002
DEFINITION wimk8.pk0025.g4 wimk8 Trifolium aestivum cDNA clone wimk8.pk0025.g4
5', end, mRNA sequence.
CA696313
CA696313.1 GI:25418099
EST.
KEYWORDS Trifolium aestivum (bread wheat)
SOURCE Trifolium aestivum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Trifoliaceae; Trifolium.
1 (bases 1 to 407)
Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,
Miao, G., Caraher, N. and Hanafey, M.K.
Dupont Wheat cDNA Sequence
Unpublished
JOURNAL Contact: Scott V. Tingey
Crop Genetics
E. I. Dupont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
location/Qualifiers
FEATURES
source
1..407
/organism="Trifolium aestivum"
/mol_type="mRNA"
/cultivar="Stephens"
/db_xref="taxon:4565"
/clone="wimk8.pk0025.g4"
/cissue_type="leaf"
/clone_lib="wimk8"
/note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
XhoI; Wheat (Trifolium aestivum L.) seedlings 8 hr after
inoculation w/ E. graminis and
6-iodo-3-propyl-2-propyloxy-4(3H)-quinazolinone"
BASE COUNT 147 a 71 c 77 g 112 t
ORIGIN

Query Match 2.3%; Score 30; DB 14; Length 407;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1294 CAAGTAAAAAATCTCGAG 1323
|||||
35 CAAGTAAAAAATCTCGAG 6

RESULT 11
LOCUS AA126428/c 447 bp mRNA linear EST 26-NOV-1996
DEFINITION zn84h09.s1 Stragene lung carcinoma 937218 Homo sapiens cDNA clone
IMAGE:564929 3', mRNA sequence.
AA126428
AA126428.1 GI:1686068
EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 447)
Hiller, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapelli, B.,
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins
M., Hultman, M., Kucaba, T., Lacey, M., Le, M., Le, N., Mardis, E., Moore
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaatis, E.,
Underwood, K., Woldmann, P., Watson, R., Wilson, R. and Wirtz, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
MEDLINE
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estowatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.jnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 359.

```

FEATURES
  source
    Location/Qualifiers
      1..447
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="GDB:4583993"
        /db_xref="taxon:9606"
        /clone="IMAGE:564929"
        /tissue_type="lung carcinoma"
        /cell_line="NCI-H69"
        /dev_stage="cell line NCI-H69"
        /lab_host="SOIR (kanamycin resistant)"
        /clone_id="Stratagene lung carcinoma 937218"
        /note="Organ: lung; Vector: pBluescript SK-; Site 1: EcoRI
; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo
dt. Small cell carcinoma cell line NCI-H69. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGCGACGAG 3' ~3' adaptor sequence: 5'
CTCAGCTTTTCTTTTCTTTT 3'."
BASE COUNT      120 a      83 c      77 g      164 t      3 others
ORIGIN
Query Match      2.3%; Score 30; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1293 TCAAGTAAAAAAAAAAAAAAAACTCGA 1322
         |||||||
         42 TCAAGTAAAAAAAAAAAAAAAACTCGA 13

RESULT 12
BU721218      515 bp      mRNA      linear      EST 12-FEB-2003
LOCUS      BU721218      SUM Schistosoma japonicum cDNA, mRNA sequence.
ACCESSION      BU721218
VERSION      BU721218.1 GI:28328565
KEYWORDS      EST.
SOURCE      Schistosoma japonicum
ORGANISM      Schistosoma japonicum
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigoidida; Schistosomatidae; Schistosomatidae; Schistosoma.
1 (bases 1 to 515)
Hu.W., Yan,Q., Shen,D., Liu,F., Xu,X., Zhu,Z., Zhang,X., Wang,J.,
Sun,J., Xu,X., Wang,Z., Zeng,L., Rong,Y., Wu,X., Qu,J., Xu,Z.,
Hang,J., Ma,Y., Wang,S., Wang,Z., Xue,C., Feng,Z., Chen,Z. and Han
Z.
Expressed sequence tags from male adults of Schistosoma japonicum
JOURNAL
  title      Unpublished
  contact      Zeguang Han
  address      Chinese National Human Genome Center at Shanghai
                351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
                201203, P. R. China
                Tel: 86-21-50801919 (ex. 45)
                Fax: 86-21-50801922
                Email: hanzg@chgc.sh.cn.
  location/Qualifiers
    1..515
      /organism="Schistosoma japonicum"
      /mol_type="mRNA"
      /db_xref="taxon:6182"
      /sex="male"
      /tissue_type="whole body"
      /dev_stage="adult"
      /lab_host="rabbits"
      /clone_id="SUM"
BASE COUNT      156 a      83 c      79 g      197 t
ORIGIN
Query Match      2.3%; Score 30; DB 13; Length 515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1294 CAACTAAAAAAAAAAAAAAAACTCGAG 1323

```

```

Db      35 CAACTAAAAAAAAAAAAAAAACTCGAG 6
|||||
RESULT 13
B0081761      559 bp      mRNA      linear      EST 04-APR-2002
LOCUS      B0081761      san28h07.y1 Gm-cl084 Glycine max cDNA clone SOYBEAN CLONE ID:
DEFINITION      Gm-cl084-5365 5' similar to TR:Q9X142 Q9X142 F9L1.21 PROTEIN. ;
                mRNA sequence.
B0081761      GI:19938525
ACCESSION      B0081761.1 GI:19938525
VERSION      EST.
KEYWORDS      Glycine max (soybean)
SOURCE      Glycine max
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
                ; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                Glycine.
                1 (bases 1 to 559)
REFERENCE      Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khana
                A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
                Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
                ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
                ,R., Rittler,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCam
                ,R., Waterston,R. and Wilson,R.
                Public Soybean EST Project
                Unpublished
                Contact: Shoemaker R/Public Soybean EST Project
                Public Soybean EST Project
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                Tel: 314 286 1800
                Fax: 314 286 1810
                Email: est@watson.wustl.edu
                This clone is available through: Resgen, Invitrogen Corp. 2130
                South Memorial Parkway Huntville, AL 35801 For further information
                call: (800)-533-4363 or contact: cdu@resgen.com web site:
                www.resgen.com
                Seq primer: -40RP from Gibco
                High quality sequence stop: 421.
FEATURES
  source
    Location/Qualifiers
      1..559
        /organism="Glycine max"
        /mol_type="mRNA"
        /db_xref="taxon:3847"
        /clone="SOYBEAN CLONE ID: Gm-cl084-5365"
        /tissue_type="Etiolated hypocotyls (Williams 82)"
        /lab_host="DH10B"
        /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed by M. Bhattacharya
from mRNA isolated from etiolated hypocotyls from the
cultivar Williams 82. Tissue was inoculated with
Phytophthora soyae race 1 and tissues were harvested 2 and
4 hours following infection. The library is the pool of
these two time points. Complementary DNA was synthesized
from mRNA using a primer consisting of a poly(dT) sequence
with a XhoI restriction site. EcoRI adapters were ligated
to the blunt-ended cDNA fragments followed by XhoI
digestion. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
DH10B host cells (Gibco BRL). This library was constructed
by M. Bhattacharya in the laboratory of Dr. Randy
Shoemaker at Iowa State University."
BASE COUNT      165 a      94 c      140 g      160 t
ORIGIN
Query Match      2.3%; Score 30; DB 13; Length 559;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1293 TCAAGTAAAAAAAAAAAAAAAACTCGA 1322
 Db 530 TCAAGTAAAAAAAAAAAAAAAACTCGA 559

RESULT 14
 AM334649/c 603 bp mRNA linear EST 31-JAN-2000
 LOCUS S38G4 AGS-1 Pneumocystis carinii cDNA 3', mRNA sequence.
 DEFINITION AM334649
 ACCESSION AM334649
 VERSION AM334649.1 GI:6831110
 KEYWORDS EST.
 SOURCE Pneumocystis carinii
 ORGANISM Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
 Pneumocystidaceae; Pneumocystis.

REFERENCE 1 (bases 1 to 603)
 AUTHORS Smullan, A.G., Arnold, J., Weise, M., Wunderlich, J., Staben, C., Edman, J.C., Kovacs, J. and Cushman, M.
 TITLE Expressed sequence tags from *Pneumocystis carinii*
 JOURNAL Unpublished
 COMMENT Contact: Stephen C
 School of Biological Sciences
 University of Kentucky
 101 Morgan Building, University of Kentucky, Lexington, KY
 40506-0225, USA
 Tel: 606 257 2161
 Fax: 606 257 1717
 Email: staben@pop.uky.edu.
 Location/Qualifiers

FEATURES
 source
 1..603
 /organism="Pneumocystis carinii"
 /mol_type="mRNA"
 /db_xref="taxon:4754"
 /lab_host="E. coli"
 /clone_lib="AGS-1"
 /note="Vector: Lambda ZAP II; Site 1: EcoRI; Site 2: XhoI;
 P. carinii organisms (3x10e9) from a single rat (59-1-6,
 sacrificed on 3/17/99) at Cincinnati VA facilities.
 Trizol extracted RNA. Oligo dt priming. Standard
 conditions described by vendor, Stratagene. Further
 details see www.uky.edu/Project/pneumocystis/"

BASE COUNT 198 a 82 c 88 g 235 t
 ORIGIN

Query Match 2.3%; Score 30; DB 9; Length 603;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1294 CAAGTAAAAAAAAAAAAAAAACTCGAG 1323
 Db 41 CAAGTAAAAAAAAAAAAAAAACTCGAG 12

RESULT 15
 BE661929 647 bp mRNA linear EST 06-SEP-2000
 LOCUS 1240 GmaxSC Glycine max cDNA, mRNA sequence.
 DEFINITION BE661929
 ACCESSION BE661929
 VERSION BE661929.1 GI:9987821
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)

REFERENCE 1 (bases 1 to 647)
 AUTHORS Harris, N., Chapman, B.P. and Gijzen, M.
 TITLE Gene expression in developing soybean seed coats
 JOURNAL Unpublished
 COMMENT Contact: Gijzen M
 Agriculture and Agri-Food Canada

1391 Sandford Street, London, Ontario, Canada N5V 4T3
 Tel: 519 457 1470
 Fax: 519 457 3997
 Email: gijzen@em.agr.ca.
 Location/Qualifiers

FEATURES
 source
 1..647
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Haro60 63"
 /db_xref="taxon:3847"
 /tissue_type="Seed coats"
 /lab_host="E. coli strain XLOLR"
 /clone_lib="GmaxSC"
 /note="Vector: pBR-CMW; Site 1: EcoRI; Site 2: XhoI; This
 cDNA library was constructed from polyA+ enriched mRNA
 from green seed coats in mid to late developmental stage
 , average fresh weight 250 mg per seed. Traces of pod and
 embryo tissue also present. Complementary DNA was
 synthesized from mRNA using an XhoI-poly(dT)
 linker-primer. EcoRI adapters were ligated to the
 blunt-ended cDNA fragments and the products were digested
 with XhoI for directional cloning into lambda ZAP Express
 vector. This lambda library was amplified once using E.
 coli host strain XL1 Blue MRF'. Inserts were then
 subcloned by mass excision using EXASist helper phage for
 conversion into phagemid vector pBR-CMW in E. coli host
 strain XLOLR."

BASE COUNT 201 a 153 c 128 g 164 t 1 others
 ORIGIN

Query Match 2.3%; Score 30; DB 10; Length 647;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1294 CAAGTAAAAAAAAAAAAAAAACTCGAG 1323
 Db 599 CAAGTAAAAAAAAAAAAAAAACTCGAG 628

RESULT 16
 BE641651/c 678 bp mRNA linear EST 19-DEC-2000
 LOCUS NF051F03IN1F030 Insect herbivory Medicago truncatula cDNA clone
 DEFINITION NF051F03IN 5', mRNA sequence.
 ACCESSION BE641651
 VERSION BE641651.1 GI:11905809
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE 1 (bases 1 to 678)
 AUTHORS Korth, K., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores
 H.R., Inman, J.T., Weller, J.W. and May, G.D.
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 JOURNAL Medicago truncatula insect herbivory library
 COMMENT Unpublished
 Contact: Korth K
 Dept. of Plant Pathology
 University of Arkansas
 217 Plant Science Building, Fayetteville, AR 72701, USA
 Tel: 501 575 5191
 Fax: 501 575 7601
 Email: kthort@comp.uark.edu
 Insert Length: 678 Std Error: 0.00
 Plate: 051 row: F column: 03
 Seq primer: TCACACAGAAACGCTATGAC.

FEATURES
 source
 1..678
 /organism="Medicago truncatula"
 /mol_type="mRNA"

/db xref="taxon:3880"
/clone="NF051F03IN"
/tissue_type="local and systemic leaves"
/dev_stage="mature"
/note="Vector: Lambda Zap; Library was produced from fully expanded M. truncatula leaves of plants fed upon by Spodoptera exigua (beet armyworm) for 24 hours. Systemic (undamaged leaves from injured plants) and wounded leaves were harvested and pooled."

BASE COUNT 265 a 114 c 97 g 200 t 2 others
ORIGIN

Query Match 2.3%; Score 30; DB 10; Length 678;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1294 CAACTAATAAAAAAAAAAACTCGAG 1323
Db 46 CAACTAATAAAAAAAAAAACTCGAG 17

RESULT 17
BF316408 867 bp mRNA linear EST 21-NOV-2000
LOCUS 601901981F1 NIH_MGC_19 Homo sapiens CDNA clone IMAGE:4130973 5',
DEFINITION mRNA sequence.

ACCESSION BF316408
VERSION BF316408.1 GI:11264765
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: image.lnl.gov
Plate: LHCMI028 row: a column: 22
High quality sequence stop: 561.
Location/Qualifiers

FEATURES
source 1..867
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4130973"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 19"
/note="Organ: brain; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 203 a 246 c 261 g 157 t
ORIGIN

Query Match 2.3%; Score 30; DB 10; Length 867;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1294 CAACTAATAAAAAAAAAAACTCGAG 1323

Db 47 CAACTAATAAAAAAAAAAACTCGAG 18

RESULT 18
BM130194 166 bp mRNA linear EST 02-NOV-2002
LOCUS BM130194 Nori Satoh unpublished CDNA library, gastrula and neurula
DEFINITION Clona intestinalis CDNA clone rcig023e01 3', mRNA sequence.
ACCESSION BM130194
VERSION BM130194.1 GI:24486593
KEYWORDS EST.
SOURCE Clona intestinalis
ORGANISM Clona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Plepobranchia; Clonidae; Clona.

REFERENCE 1 (bases 1 to 166)
AUTHORS Satou, Y., Shih-I, T., Kohara, Y. and Satoh, N.
TITLES Expressed genes in Clona intestinalis (2002c)
JOURNAL Unpublished
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@acidiaf.zool.kyoto-u.ac.jp.
Location/Qualifiers

FEATURES
source 1..166
/organism="Clona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcig023e01"
/tissue_type="whole body"
/dev_stage="gastrula and neurula"
/clone_lib="Nori Satoh unpublished CDNA library, gastrula and neurula"

BASE COUNT 31 a 49 c 18 g 65 t 3 others
ORIGIN

Query Match 2.2%; Score 29; DB 13; Length 166;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
Db 150 AAGTAAAAAAAAAAAAAAAAAACTCGAG 122

RESULT 19
CA671233 259 bp mRNA linear EST 24-NOV-2002
LOCUS wlsu2.pk0007.d10 wlsu2 Triticum aestivum CDNA clone
DEFINITION wlsu2.pk0007.d10 5' end, mRNA sequence.
ACCESSION CA671233
VERSION CA671233.1 GI:25249947
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
/ Triticaceae; Triticum.
1 (bases 1 to 259)
Tingey, S. V., Powell, W., Wolters, P., Dolan, M., Hatney, C., Yuan, Z., Miao, G., Caraher, N. and Hanafey, M.K.
Dupont Wheat cDNA Sequence
Unpublished
Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607

REFERENCE 1 (bases 1 to 259)
AUTHORS Tingey, S. V., Powell, W., Wolters, P., Dolan, M., Hatney, C., Yuan, Z., Miao, G., Caraher, N. and Hanafey, M.K.
TITLES Dupont Wheat cDNA Sequence
JOURNAL Unpublished
COMMENT Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607

Query Match 2.3%; Score 30; DB 10; Length 867;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1294 CAACTAATAAAAAAAAAAACTCGAG 1323

Email: Scott.V.Tingey@usa.dupont.com
Seq primer: M13.

FEATURES
source Location/Qualifiers

1. 259
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Stephens"
/db_xref="taxon:4565"
/clone="w18u2.pk007.d10"
/issue_type="leaf"
/clone_id="w18u2"
/note="Vector: pGEM-T Easy; Site 1: EcoRI; Site 2: XhoI;
wheat (Triticum aestivum L.) w18 cDNA subcloned with
w180 cDNA"

BASE COUNT 76 a 67 c 58 g 58 t

ORIGIN

Query Match 2.2%; Score 29; DB 14; Length 259;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1295 AAGTAAAAAAAAAACTCGAG 1323
|||||
41 AAGTAAAAAAAAAACTCGAG 13

RESULT 20
CA729647/c 261 bp mRNA linear EST 26-NOV-2002

LOCUS w1pic.pk001.d10 w1pic Triticum aestivum cDNA clone w1pic.pk001.d10
DEFINITION 5' end, mRNA sequence.

ACCESSION CA729647
VERSION CA729647.1 GI:25451650

KEYWORDS EST
SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Triticum.

1 (bases 1 to 261)
Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,
Miao, G., Caraher, N. and Hanafey, M.K.

Unpublished
Contact: Scott V. Tingey
Crop Genetics

E. I. Dupont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607

Email: Scott.V.Tingey@usa.dupont.com
Seq primer: M13.

FEATURES
source Location/Qualifiers

1. 261
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="w1pic.pk001.d10"
/issue_type="leaf"
/lab_host="DH10B"
/clone_id="w1pic"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
XhoI; wheat (Triticum aestivum, Hi Line) immature plants"

BASE COUNT 79 a 67 c 55 g 55 t

ORIGIN

Query Match 2.2%; Score 29; DB 14; Length 261;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1295 AAGTAAAAAAAAAACTCGAG 1323
|||||
35 AAGTAAAAAAAAAACTCGAG 7

RESULT 21
BG232362/c 269 bp mRNA linear EST 10-MAY-2001

LOCUS pat0a04.y1 Hawdon Ancylostoma caninum L3 Ancylostoma caninum cDNA

DEFINITION 5', mRNA sequence.
ACCESSION BG232362
VERSION BG232362.1 GI:12727536

KEYWORDS EST
SOURCE Ancylostoma caninum (dog hookworm)

ORGANISM Ancylostoma caninum
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.

1 (bases 1 to 269)
McCarter, J., Clifton, S., Chapelli, B., Pape, D., Martin, J., Wylie, T.,
Dante, M., Marrs, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.,
Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, R.,
Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe,
M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S.,
Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and
Wilson, R.

The Washington Univ. Nematode EST Project, 1999
Unpublished
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
The library was constructed by Dr. John Hawdon of Yale University
(john.hawdon@yale.edu). DNA Sequencing by: Washington University
Genome Sequencing Center, St. Louis.

Seq primer: 40RP from Gibco
High quality sequence scop: 249.

FEATURES
source Location/Qualifiers

1. 269
/organism="Ancylostoma caninum"
/mol_type="mRNA"
/strain="Shanghai strain"
/db_xref="taxon:29170"
/sex="female and male"
/dev_stage="3rd stage larva (L3)"
/lab_host="E. coli, XL-1 Blue MRF (Stratagene)"
/clone_id="Hawdon Ancylostoma caninum L3"
/note="Vector: Lambda Uni-ZAP XR Vector (Stratagene);
Site 1: EcoRI (5'); Site 2: XhoI (3'); Oligo (dT) primed
library from L3 worms isolated from dog. cDNA was
constructed and cloned unidirectionally into the vector
from the EcoRI to the XhoI site. The library went
through one round of amplification."

BASE COUNT 96 a 45 c 58 g 70 t

ORIGIN

Query Match 2.2%; Score 29; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1295 AAGTAAAAAAAAAACTCGAG 1323
|||||
34 AAGTAAAAAAAAAACTCGAG 6

RESULT 22
AA325025/c 294 bp mRNA linear EST 20-APR-1997

LOCUS EST7949 Cerebellum II Homo sapiens cDNA 5' end similar to EST

DEFINITION containing L1 repeat, mRNA sequence.

ACCESSION AA325025
VERSION AA325025.1 GI:1977271

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 294)
AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult
C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White
O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C., Claydon, R.A.,
Cline, T.R., Coston, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald,
L.M., Fitzhugh, W.M., Fritchman, J.L., Geophagen, N.S., Giodok, A.,
Ghem, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,
Moreno-Balaguer, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.U., Sauder, D.W., Shirley, R.,
Small, K.V., Spriggs, T.A., Utechtack, T.R., Weidman, J.F., Li, Y.,
Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A., He, W.W.,
Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,
Kunisch, C., Hungjun, J., Li, H., Weisner, P.S., Olsen, H., Raymond, L.,
Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon
M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and
Venter, J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhoc):125520"
/db_xref="taxon:9606"
/tissue_type="cerebellum"
/dev stage="adult"
/clone_lib="Cerebellum II"
/note="Organ: brain; Vector: pbluescript SK-; Site_1:
ECORI; Site_2: XhoI"
BASE COUNT 68 a 60 c 56 g 110 t
ORIGIN
Query Match 2.2%; Score 29; DB 9; Length 294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1295 AAGTAAAAAAAAAAAAAAAACTCGAG 1323
Db 47 AAGTAAAAAAAAAAAAAAAACTCGAG 19

REFERENCE 1 (bases 1 to 301)
AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
K., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood
K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E.,
Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
and Wilson, R.
TITLE Mashu zebrafish EST Project 1998
JOURNAL Unpublished
COMMENT Other ESTs: ft30b12.x1
Contact: Stephen L. Johnson
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbratfish@watson.wustl.edu
The library was constructed by Dr. Z. Gong. DNA Sequencing by:
Washington University Genome Sequencing Center St. Louis. Please
contact Zhiyuan Gong for further information on this library
(National University of Singapore, Department of Biological
Sciences, Lower Kent Ridge Road, Singapore 119260).
zebrafish identity (p-value greater than 1e-99) found to:
g1|2446831|gb|AA606178|AA606178 fa28d11.61 Ekkerpost segmentation
zebrafish
Seq primer: T3 ET from Amerham.
Location/Qualifiers
1. 301
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:5152079"
/sex="male"
/dev stage="4-5 month"
/lab host="DH10B (phage-resistant)"
/note="Organ: testis (pooled); Vector: pbluescript SK-;
Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolated from
the testes of 31 male adult zebrafish (4-5 month old).
cDNAs were made using oligo-dT primers and inserted into
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
mass-excisid to pbluescript SK- following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Zhiyuan Gong for further information on
this library (National University of Singapore,
Department of Biological Sciences, Lower Kent Ridge Road,
Singapore 119260)."
BASE COUNT 90 a 54 c 76 g 81 t
ORIGIN
Query Match 2.2%; Score 29; DB 12; Length 301;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1295 AAGTAAAAAAAAAAAAAAAACTCGAG 1323
Db 272 AAGTAAAAAAAAAAAAAAAACTCGAG 300

RESULT 24
LOCUS BM530269
DEFINITION BM530269 332 bp mRNA linear EST 19-FEB-2002
IMAGE:561606 5' similar to SW-DUT RAT P70583 DEXOXYRIDINE
5'-TRIPHOSPHATE NUCLEOTIDOHYDROLASE ; mRNA sequence.
ACCESSION BM530269
VERSION BM530269.1 GI:18738032
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

1 Cyprinidae; Danilo.
1 (bases 1 to 332)
Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
S., Hillier, L., Kueba, T., Martin, J., Beck, C., Wyllie, T., Underwood
K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritzer, E.,
Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
and Wilson, R.
Mashu zebrafish EST Project 1998
Unpublished
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrfish@wustl.edu
The library was constructed by Dr. Z. Gong. DNA Sequencing by:
Washington University Genome Sequencing Center St. Louis. Please
contact Zhiyuan Gong for further information on this library
(National University of Singapore, Department of Biological
Sciences, Lower Kent Ridge Road, Singapore 119260).
Seq primer: T3 ET from Amersham.
Location/Qualifiers
1. .332
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:561606"
/sex="male"
/dev_stage="4-5 month"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Gong zebrafish testis"
/note="Organ: testis (pooled); Vector: pBluescript SK-;
site_1: XhoI; site_2: EcoRI; Poly A: RNA was isolated from
the testes of 31 male adult zebrafish (4-5 month old).
cDNAs were made using oligo-dT primers and inserted into
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
messenger-extracted to pBluescript SK- following the Washington
University protocol
(http://genome.wustl.edu/sec/lambda_protocol.shtml).
Please contact Zhiyuan Gong for further information on
this library (National University of Singapore,
Department of Biological Sciences, Lower Kent Ridge Road,
Singapore 119260)."

BASE COUNT
94 a 54 c 90 g 94 t

ORIGIN

Query Match 2.2%; Score 29; DB 12; Length 332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1295 AAGTAAAAAAAAAAAAAAAAAATCGAG 1323
|||||
Db 302 AAGTAAAAAAAAAAAAAAAAAATCGAG 330

RESULT 25
BI233432 364 bp mRNA linear EST 07-SEP-2001
LOCUS 949078B02.y1 949 - Juvenile leaf and shoot cDNA from Steve Moose
DEFINITION Zea mays cDNA, mRNA sequence.
ACCESSION BI233432
VERSION BI233432.1 GI:15499532
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 364)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY

JOURNAL
COMMENT
Unpublished
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 949076 row: B column: 02.
Location/Qualifiers
1. .364
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W64A"
/db_xref="taxon:4577"
/tissue_type="Immature leaf primordium and vegetative
meristem"
/dev_stage="4 stages from 3-13 days after imbibing"
/lab_host="E. coli XL0LR"
/clone_lib="949 - Juvenile leaf and shoot cDNA from Steve
Moose."

BASE COUNT
102 a 67 c 98 g 97 t

ORIGIN

Query Match 2.2%; Score 29; DB 12; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1295 AAGTAAAAAAAAAAAAAAAAAATCGAG 1323
|||||
Db 331 AAGTAAAAAAAAAAAAAAAAAATCGAG 359

RESULT 26
BI233432 367 bp mRNA linear EST 11-JUL-2001
LOCUS 949010A01.y2 949 - Juvenile leaf and shoot cDNA from Steve Moose
DEFINITION Zea mays cDNA, mRNA sequence.
ACCESSION BI233432
VERSION BI233432.1 GI:14701014
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 367)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
COMMENT Unpublished
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 949010 row: A column: 01.
Location/Qualifiers

source

1. .367
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W64A"
/db_xref="taxon:4577"
/tissue_type="immature leaf primordium and vegetative meristem"
/dev_stage="4 stages from 3-13 days after imbibing"
/lab_host="E. coli XL0LR"
/clone_lib="949 - Juvenile leaf and shoot cDNA from Steve Moose"

/note="Organ: juvenile vegetative shoots; Vector: pAD-GAL4-2.1; Site 1: EORI; Site 2: XhoI; Equal amounts of total RNA by weight from 4 tissue sources (see below) were pooled, polyA+ RNA isolated, and cDNA synthesized for EcoRI (5') and XhoI (3') directional cloning into lambda Hyrizap vector from Stratagene. Tissue Sources: 1. Whole shoots 3 days after sowing/imbibing in wet soil. 2. Basal 1.5 cm shoots 6 days after sowing - includes yellow portions of developing leaves 1-5, primordia from 6-8, and the vegetative apex. 3. Non-green portions of developing leaves 4-5 and the vegetative apex, including adult leaf primordia, 9 days after sowing. 4. Partially expanded and greening leaves 4-5 at 13 days after sowing."

BASE COUNT 103 a 66 c 99 g 99 t

ORIGIN

Query Match 2.2%; Score 29; DB 12; Length 367;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
|||
Db 334 AAGTAAAAAAAAAAAAAAAAAACTCGAG 362

RESULT 27
BE445989 377 bp mRNA linear EST 25-JUN-2000
LOCUS WHE1148_E10_J20ZS wheat etiolated seedling root normalized cDNA
DEFINITION library Triticum aestivum cDNA clone WHE1148_E10_J20, mRNA
sequence.
ACCESSION BE445989 GI:9445551
VERSION BE445989.1
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
1 (bases 1 to 377)
AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han ,P.S., Hala,C.C., Kang,Y., Lazo,G.R., Miller,R., Nguyen,H.T., Rausch,C.J., Seaton,C.B., Tong,J.C. and Zhang,D.
TITLE The structure and function of the expressed portion of the wheat genomes - Normalized root cDNA library
JOURNAL unpublished
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel.: 5105595773
Fax: 5105595818
Email: oanderson@wp.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES
source
1. .377
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"

/clone="WHE1148_E10_J20"
/tissue_type="root"
/dev_stage="Five day old etiolated seedling"
/lab_host="E. coli DH10B"
/clone_lib="wheat etiolated seedling root normalized cDNA library"
/note="Vector: lambda Uni-ZAP XR, excised phagemid pBluescript SK; Site 1: EORI; Site 2: XhoI; Seeds were surface-sterilized, germinated and grown aseptically in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Roots were harvested. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made in the T7 Close lab (Choi, Close, Fenton) at the University of California, Riverside. The cDNA clones were in vivo excised to give phagescript phagemids before normalization was carried out. The mass excision of phagemid library and normalization were done in HT Nguyen lab by D. Zhang at Texas Tech University. Normalization protocol used was that of Soares. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 96 a 107 c 92 g 82 t

ORIGIN

Query Match 2.2%; Score 29; DB 10; Length 377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
|||
Db 333 AAGTAAAAAAAAAAAAAAAAAACTCGAG 361

RESULT 28
BI674867 407 bp mRNA linear EST 12-SEP-2001
LOCUS 949076B02.Y2 949 - Juvenile leaf and shoot cDNA from Steve Moose
DEFINITION Zea mays cDNA, mRNA sequence.
ACCESSION BI674867
VERSION BI674867.1 GI:15590251
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
1 (bases 1 to 407)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 949076 row: B column: 02.
FEATURES
source
1. .407
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W64A"
/db_xref="taxon:4577"
/tissue_type="immature leaf primordium and vegetative meristem"
/dev_stage="4 stages from 3-13 days after imbibing"
/lab_host="E. coli XL0LR"
/clone_lib="949 - Juvenile leaf and shoot cDNA from Steve Moose"
/note="Organ: juvenile vegetative shoots; Vector:

National Institute of Agrobiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmits@nias.affrc.go.jp
method:uni-directional, sequence direction:sequenced from T3 primer
(5' -> 3').

FEATURES
source
1. 516
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="epv30368"
/issue_type="epidermis"
/dev_stage="5th instar larva day 3"
/clone_1ib="epv3"

BASE COUNT 160 a 105 c 125 g 122 t
ORIGIN

Query Match 2.2%; Score 29; DB 12; Length 516;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1295 AAGTAAAAAAAAAACTCGAG 1323
|||||
Db 451 AAGTAAAAAAAAAACTCGAG 479

RESULT 35
LOCUS BP123880 516 bp mRNA linear EST 16-MAY-2003
DEFINITION BP123880 epv3 Bombyx mori cDNA clone epv31585, mRNA sequence.
ACCESSION BP123880
VERSION BP123880.1 GI:29556922
KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)
ORGANISM Bombyx mori

REFERENCE
AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.
1 (bases 1 to 516)
Mita,K., Moriyio,M., Shimada,T., Okano,K. and Maeda,S.
Bombyx mori cDNA (Mita,K. 2003)
Unpublished

COMMENT
Contact: Mita K
Genome Research Group
National Institute of Agrobiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmits@nias.affrc.go.jp
method:uni-directional, sequence direction:sequenced from T3 primer
(5' -> 3').

FEATURES
source
1. 516
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="epv31585"
/issue_type="epidermis"
/dev_stage="5th instar larva day 3"
/clone_1ib="epv3"

BASE COUNT 158 a 130 c 112 g 115 t 1 others
ORIGIN

Query Match 2.2%; Score 29; DB 12; Length 516;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1295 AAGTAAAAAAAAAACTCGAG 1323
|||||
Db 470 AAGTAAAAAAAAAACTCGAG 498

RESULT 36
LOCUS BP123966 516 bp mRNA linear EST 16-MAY-2003

DEFINITION BP123966 epv3 Bombyx mori cDNA clone epv31721, mRNA sequence.
ACCESSION BP123966
VERSION BP123966.1 GI:29557008
KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)
ORGANISM Bombyx mori

REFERENCE
AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.
1 (bases 1 to 516)
Mita,K., Moriyio,M., Shimada,T., Okano,K. and Maeda,S.
Bombyx mori cDNA (Mita,K. 2003)
Unpublished

COMMENT
Contact: Mita K
Genome Research Group
National Institute of Agrobiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmits@nias.affrc.go.jp
method:uni-directional, sequence direction:sequenced from T3 primer
(5' -> 3').

FEATURES
source
1. 516
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="epv31721"
/issue_type="epidermis"
/dev_stage="5th instar larva day 3"
/clone_1ib="epv3"

BASE COUNT 159 a 129 c 111 g 117 t
ORIGIN

Query Match 2.2%; Score 29; DB 12; Length 516;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1295 AAGTAAAAAAAAAACTCGAG 1323
|||||
Db 481 AAGTAAAAAAAAAACTCGAG 509

RESULT 37
LOCUS BP124078 516 bp mRNA linear EST 16-MAY-2003
DEFINITION BP124078 epv3 Bombyx mori cDNA clone epv31877, mRNA sequence.
ACCESSION BP124078
VERSION BP124078.1 GI:29557120
KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)
ORGANISM Bombyx mori

REFERENCE
AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.
1 (bases 1 to 516)
Mita,K., Moriyio,M., Shimada,T., Okano,K. and Maeda,S.
Bombyx mori cDNA (Mita,K. 2003)
Unpublished

COMMENT
Contact: Mita K
Genome Research Group
National Institute of Agrobiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmits@nias.affrc.go.jp
method:uni-directional, sequence direction:sequenced from T3 primer
(5' -> 3').

FEATURES
source
1. 516
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="epv31877"
/issue_type="epidermis"
/dev_stage="5th instar larva day 3"
/clone_1ib="epv3"

BASE COUNT 159 a 130 c 112 g 115 t
 ORIGIN

Query Match 2.2%; Score 29; DB 12; Length 516;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1295 AAGTAAAAAAAAAAAAAAAACTCGAG 1323
 Db 473 AAGTAAAAAAAAAAAAAAAACTCGAG 501

RESULT 38
 BP124196 516 bp mRNA linear EST 16-MAY-2003
 LOCUS BP124196 epv3 Bombyx mori cDNA clone epv32014, mRNA sequence.
 DEFINITION
 ACCSSION BP124196
 VERSION BP124196.1 GI:29557238
 KEYWORDS
 SOURCE Bombyx mori (domestic silkworm)
 ORGANISM Bombyx mori

REFERENCE
 AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Bombycoidea; Bombycidae; Bombyx.
 1 (bases 1 to 516)
 Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.
 Bombyx mori cDNA (Mita,K. 2003)
 TITLE Unpublished
 JOURNAL Contact: Mita K

COMMENT
 Genome Research Group
 National Institute of Agrobiological Sciences
 Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
 Email: kmits@nias.affrc.go.jp
 method:uni-directional, sequence direction:sequenced from T3 primer
 (5' -> 3').

FEATURES
 source Location/Qualifiers
 1..516
 /organism="Bombyx mori"
 /mol_type="mRNA"
 /db_xref="taxon:7091"
 /clone="epv32014"
 /tissue_type="epidermis"
 /dev_stage="5th instar larva day 3"
 /clone_1lb="epv3"

BASE COUNT 159 a 130 c 112 g 115 t
 ORIGIN

Query Match 2.2%; Score 29; DB 12; Length 516;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1295 AAGTAAAAAAAAAAAAAAAACTCGAG 1323
 Db 470 AAGTAAAAAAAAAAAAAAAACTCGAG 498

RESULT 39
 BP124420 516 bp mRNA linear EST 16-MAY-2003
 LOCUS BP124420 epv3 Bombyx mori cDNA clone epv32344, mRNA sequence.
 DEFINITION
 ACCSSION BP124420
 VERSION BP124420.1 GI:29557462
 KEYWORDS
 SOURCE Bombyx mori (domestic silkworm)
 ORGANISM Bombyx mori

REFERENCE
 AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Bombycoidea; Bombycidae; Bombyx.
 1 (bases 1 to 516)
 Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.
 Bombyx mori cDNA (Mita,K. 2003)
 TITLE Unpublished
 JOURNAL Contact: Mita K

COMMENT
 Genome Research Group
 National Institute of Agrobiological Sciences
 Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
 Email: kmits@nias.affrc.go.jp
 method:uni-directional, sequence direction:sequenced from T3 primer
 (5' -> 3').

BASE COUNT 157 a 130 c 112 g 117 t
 ORIGIN

Query Match 2.2%; Score 29; DB 12; Length 516;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1295 AAGTAAAAAAAAAAAAAAAACTCGAG 1323
 Db 469 AAGTAAAAAAAAAAAAAAAACTCGAG 497

RESULT 40
 BM303636 518 bp mRNA linear EST 02-JAN-2002
 LOCUS SNEST4a62d10.y1 csn 1 S neuropa invitro mezozoite cDNA Sarcocystis
 DEFINITION neuropa cDNA 5', mRNA sequence.
 ACCSSION BM303636
 VERSION BM303636.1 GI:18035340
 KEYWORDS
 SOURCE EST.
 ORGANISM Sarcocystis neuropa
 Sarcocystis neuropa
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Sarcocystidae; Sarcocystis.

REFERENCE
 AUTHORS 1 (bases 1 to 518)
 Howe,D.K., Stamper,S., Tang,K., Sibley,L.D., Clifton,S., Marra,M.,
 Hillier,L., Pape,D., Martin,J., Wylie,T., Theising,B., Bowers,Y.,
 Gibbons,M., Ritzer,E., McCann,R., Bissett,A., Bennett,J., Schmitt
 A., Ronko,I., Tsagareishvili,R., Fedele,M., Belaygorod,L.,
 Franklin,C., Carr,L.M., Grow,A., Maguire,L., Watkins,J., Ritchey,D.,
 Waterston,R. and Wilson,R.
 Sarcocystis neuropa EST project
 Unpublished
 Contact: Daniel K. Howe
 Sarcocystis neuropa EST project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Contact Daniel K. Howe (dkhowe2@pop.uky.edu) for further
 information relating to organism, libraries, or clone availability.
 Seq primer: -40RP from Gibco
 High quality sequence stop: 422.

FEATURES
 source Location/Qualifiers
 1..518
 /organism="Sarcocystis neuropa"
 /mol_type="mRNA"
 /strain="Sn3"
 /db_xref="taxon:42890"
 /dev_stage="mezozoite"
 /lab_host="DH10B"
 /clone_1lb="csn 1 S neuropa invitro mezozoite cDNA"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI. The library was constructed by Dan Howe, University
 of Kentucky. cDNAs were synthesized from poly(A)+ RNA
 by oligo d(T) priming and directionally cloned into the
 uni-ZAP XR lambda vector. The library was mass excised


```

/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="pav30300"
/tissue_type="posterior silkland"
/dev_stage="5th instar larva day 3"
/clone_lib="pav3"

BASE COUNT      176 a      78 c      74 g      177 t      20 others
ORIGIN

Query Match      2.2%; Score 29; DB 12; Length 525;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1295 AAGTAAAAAAAAAAAAAAAACTCGAG 1323
Db      86 AAGTAAAAAAAAAAAAAAAACTCGAG 58

RESULT 44
BP123680      528 bp      mRNA      linear      EST 16-MAY-2003
LOCUS
DEFINITION    BP123680 epv3 Bombyx mori cDNA clone epv31360, mRNA sequence.
ACCESSION
VERSION       BP123680.1 GI:29556722
KEYWORDS
SOURCE        Bombyx mori (domestic silkworm)
ORGANISM      Bombyx mori

REFERENCE
AUTHORS       Mita,K., Moriyu,M., Shimada,T., Okano,K. and Maeda,S.
TITLE         Bombyx mori cDNA (Mita,K. 2003)
JOURNAL       Unpublished
COMMENT       Contact: Mita K
               Genome Research Group
               National Institute of Agrobiological Sciences
               Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
               Email: kmita@nias.affrc.go.jp
               method: uni-directional, sequence direction: sequenced from T3 primer
               (5' -> 3').

FEATURES
source
Location/Qualifiers
1..528
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="epv31360"
/tissue_type="epidermis"
/dev_stage="5th instar larva day 3"
/clone_lib="epv3"

BASE COUNT      165 a      111 c      128 g      124 t

ORIGIN

Query Match      2.2%; Score 29; DB 12; Length 528;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1295 AAGTAAAAAAAAAAAAAAAACTCGAG 1323
Db      462 AAGTAAAAAAAAAAAAAAAACTCGAG 490

RESULT 45
BP123416      544 bp      mRNA      linear      EST 16-MAY-2003
LOCUS
DEFINITION    BP123416 epv3 Bombyx mori cDNA clone epv31026, mRNA sequence.
ACCESSION
VERSION       BP123416.1 GI:29556458
KEYWORDS
SOURCE        Bombyx mori (domestic silkworm)
ORGANISM      Bombyx mori
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

```

Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.
1 (bases 1 to 544)
Mita,K., Moriyu,M., Shimada,T., Okano,K. and Maeda,S.
TITLE         Bombyx mori cDNA (Mita,K. 2003)
JOURNAL       Unpublished
COMMENT       Contact: Mita K
               Genome Research Group
               National Institute of Agrobiological Sciences
               Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
               Email: kmita@nias.affrc.go.jp
               method: uni-directional, sequence direction: sequenced from T3 primer
               (5' -> 3').

FEATURES
source
Location/Qualifiers
1..544
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="epv31026"
/tissue_type="epidermis"
/dev_stage="5th instar larva day 3"
/clone_lib="epv3"

BASE COUNT      165 a      139 c      117 g      123 t

ORIGIN

Query Match      2.2%; Score 29; DB 12; Length 544;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1295 AAGTAAAAAAAAAAAAAAAACTCGAG 1323
Db      478 AAGTAAAAAAAAAAAAAAAACTCGAG 506

RESULT 46
AI295021      577 bp      mRNA      linear      EST 19-APR-2001
LOCUS
DEFINITION    AI295021 5prime LP Drosophila melanogaster larval-early pupal P072
               Drosophila melanogaster cDNA clone LP08533 5prime similar to
               K02314: D. melanogaster metallochionein mRNA, complete cds, mRNA
               sequence.
               AI295021
               AI295021.1 GI:3944428
               EST.
               Drosophila melanogaster (fruit fly)
SOURCE        Drosophila melanogaster
ORGANISM      Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 577)
REFERENCE     Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
               Lewis,S. and Rubin,G.M.
AUTHORS       BDGP/HMI Drosophila EST Project
TITLE         Unpublished
JOURNAL       Contact: Stapleton, M.
COMMENT       BDGP
               Lawrence Berkeley National Lab
               One Cyclotron Rd, Berkeley, CA 94720, USA
               Fax: 510 486 6798
               Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
               Plate: 85 row: C column: 9
               High quality sequence stop: 511.

FEATURES
source
Location/Qualifiers
1..577
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="LP08533"
/sex="male and female"
/dev_stage="larvae-pupae"
/lab_host="DHS-alpha"
/clone_lib="LP Drosophila melanogaster larval-early pupal
P072"

```



```

/organism="Glycine max"
/mol_type="mrna"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1018-1336"
/tissue_type="leaves of greenhouse grown plants"
/dev_stage="2-3 weeks old"
/lab_host="DH10B (Gibco BRU)"
/clone_id="Gm-c1018"
/notes="Vector: Bluescript II XR; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from 2-3 week old greenhouse grown plants. The cDNA library
was prepared using the Stratagene Bluescript II XR
library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
Bluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (Gibco BRU). This
library was constructed by Dr. Randy Shoemaker and Dr.
John Eppelding."

```

BASE COUNT 173 a 144 c 113 g 171 t 3 others

ORIGIN

Query Match 2.2%; Score 29; DB 9; Length 604;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
 |||||
 33 AAGTAAAAAAAAAAAAAAAAAACTCGAG 5

RESULT 50
 AM949436
 LOCUS EST361506 617 bp mRNA linear EST 01-JUN-2000
 DEFINITION MAGE resequences, MAGA Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM949436
 VERSION AM949436.1 GI:8139063
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 617)
 Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt
 I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
 Quackenbush, J.
 Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 unpublished

JOURNAL Contact: John Quackenbush
 COMMENT The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 2

FEATURES
 source
 1. .617
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mrna"
 /db_xref="taxon:9606"
 /clone_id="MAGE resequences, MAGA"
 /note="Vector: BluescriptSKm"

BASE COUNT 175 a 129 c 87 g 226 t

ORIGIN

Query Match 2.2%; Score 29; DB 10; Length 617;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1295 AAGTAAAAAAAAAAAAAAAAAACTCGAG 1323
 |||||
 DB 513 AAGTAAAAAAAAAAAAAAAAAACTCGAG 541

Search completed: January 28, 2004, 18:12:11
 Job time : 2698 secs